

# Absolute scaling of sensory magnitudes: A validation

J. J. ZWISLOCKI and D. A. GOODMAN

*Institute for Sensory Research, Syracuse University, Syracuse, New York 13210*

The hypothesis that, in magnitude estimation and production, subjects tend to pair numbers with sensation magnitudes on absolute rather than ratio scales is tested experimentally. This implies that not only sensations, but also numbers acquire absolute psychological magnitudes. The specific experiments are performed on loudness and line lengths. The latter reveal that the subjective magnitudes of numbers are formed before the age of 6 and do not change after that age. It is suggested that the absolute coupling of numbers with sensation magnitudes originates from the concept of numerosity where numbers have absolute meanings.

This article concerns the methods of magnitude estimation and production worked out systematically by S. S. Stevens and his co-workers (e.g., Reynolds & Stevens, 1960; Stevens, 1955, 1956, 1975) and, more specifically, their modification by Hellman and Zwislöcki (1961, 1963, 1964, 1968) based on experimental evidence suggesting that subjects tend to use absolute rather than ratio scales. To be sure, a ratio scale allows multiplication by a constant (e.g., Stevens, 1951), whereas the absolute scale implies a fixed unit and, therefore, an absolute coupling between numerals and psychological magnitudes. Such a coupling is intuitively difficult to accept (Krantz, 1972) and would hardly be considered were it not for extensive empirical evidence suggesting it. Since this evidence was largely overlooked in psychological literature, its brief review appears warranted. Three examples, all showing convergence of magnitude estimation and production data obtained under diverse conditions toward one scale, should set the stage for the more systematic tests of the current study.

The first example deals with two similar experiments on magnitude estimation and production conducted in different laboratories on different groups of subjects by different teams of investigators who were not in communication with each other. The stimuli consisted of repeated 1,000-Hz tone bursts presented monaurally, and the subjects were instructed to couple numbers with loudness so that the loudness magnitude appeared equal to the subjective number magnitude. No reference standards or moduli were given so that each subject was free to choose his or her own implied unit. The medians reproduced in Figure 1 were obtained without any normalization of the individual data, so both the slope and the position of

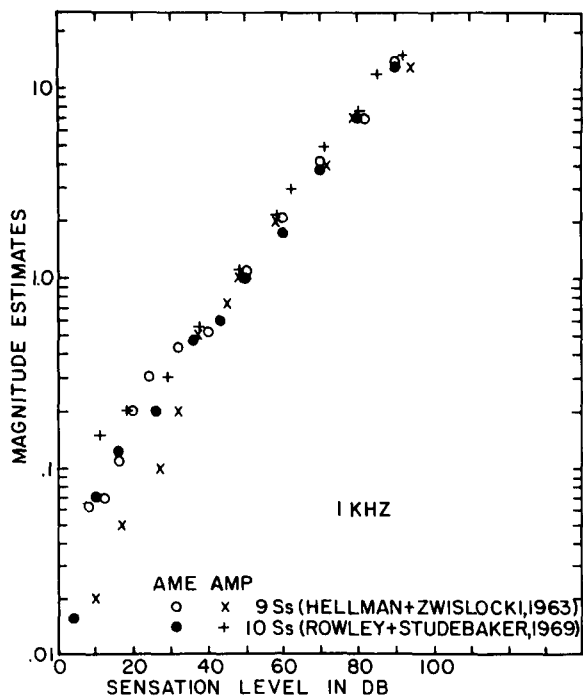


Figure 1. Absolute magnitude-estimation (AME) and absolute magnitude-production (AMP) results for loudness at 1,000 Hz obtained in two independent studies on experienced subjects (no designated standards). The points represent medians of unnormalized raw scores—the circles, for AME, the crosses for AMP. (Data from Hellman & Zwislöcki, 1963; Rowley & Studebaker, 1969.)

the curves implied by the data points are significant and represent absolute scales (AME for absolute magnitude estimation and AMP for absolute magnitude production). It is evident that, on the average, both groups associated approximately the same numbers with the same sensation levels and, presumably, the same loudnesses. This is particularly evident above the number .5 associated with a sensation level of approximately 40 dB. Lower numbers show greater variability—a common finding.

We thank Manlius Pebble Hill School for letting us use their entire first grade in our experiments and Dr. S. C. Chamberlain and Mrs. Peter Black for making the arrangements. This work was supported by NIH Grant NS-03950.

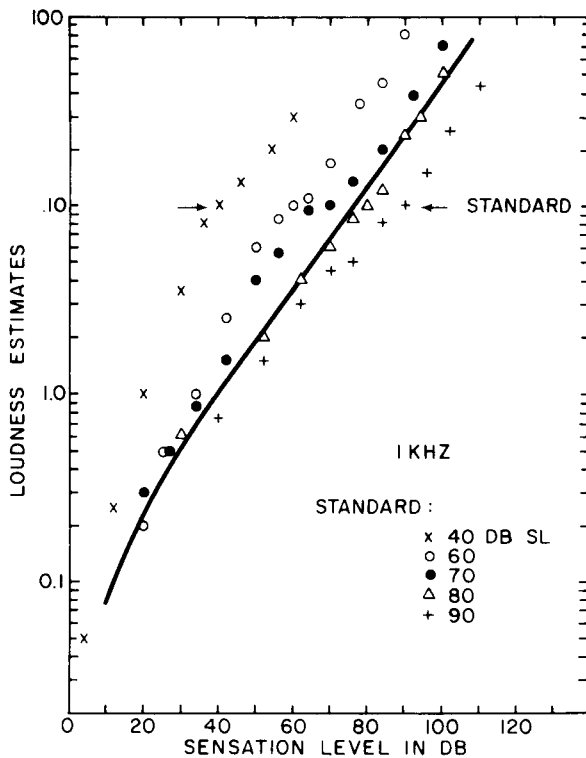


Figure 2. Median loudness estimates at 1,000 Hz obtained with the reference number 10 associated with five different reference SLs. The solid line approximates the data of Figure 1 corrected for binaural summation and agrees with the results of several independent studies. (Data from Hellman & Zwislowski, 1961.)

The remaining two examples concern magnitude estimation relative to moduli imposed by the experimenters and requiring the subjects to respond on ratio scales. The stimuli were the same as in the first example but were presented binaurally. Binaural loudness summation, a well-documented effect (e.g., Hellman & Zwislowski, 1963; Marks, 1978; Reynolds & Stevens, 1960; Scharf & Fishken, 1970), should lead to an approximate doubling of number values. Median data obtained on nine subjects with five different reference sensation levels associated with the reference number 10 are reproduced in Figure 2 (Hellman & Zwislowski, 1961). The solid line approximates the data of Figure 1 corrected for binaural summation. Above the modulus, the data points appear to follow nearly parallel curves, consistent with ratio scaling. However, below the modulus a greater range of intensities was available, and the data points converge gradually on the curve derived from the absolute-scaling procedure. The convergence is particularly clear when the number ratios become very small. A similar effect can be seen in Figure 3, whose data points are reproduced from an earlier article of Stevens (1956). A group of 16 subjects were instructed to estimate loudness ratios relative, first, to the reference number 1 associated with a

sound-pressure level of 30 dB and, then, to the reference number 100 associated with a sound-pressure level of 120 dB. The curves in the figure parallel the absolute scale of Figure 1. The solid ones are drawn through the modulus points; the dashed one joins the points far removed from the moduli. Near the moduli, the data points lie approximately on the solid lines and, therefore, are consistent with ratio scales. However, farther away, the points converge on the dashed curve irrespective of the standard. That curve coincides almost exactly with the absolute scale of Figure 1 when corrections for binaural listening and a transition from sensation level to sound-pressure level are made.

The data of Figures 2 and 3 agree in suggesting that, on the average, subjects are able to follow ratio scales in the magnitude estimation procedure within a restricted range about the modulus. Outside this range, their responses tend to converge on the function that is obtained without a designated modulus and data normalization (Figure 1), that is, on what may be interpreted as a natural absolute scale.

The hypothesis that people tend to scale sensation magnitudes on absolute scales was tested in many experiments in which a direct match of loudness (Hellman, 1976; Hellman & Zwislowski, 1963, 1964, 1968) or of subjective vibrotactile intensity (Verrillo, Fraioli, & Smith, 1969) were successfully predicted from combined absolute magnitude estimation and

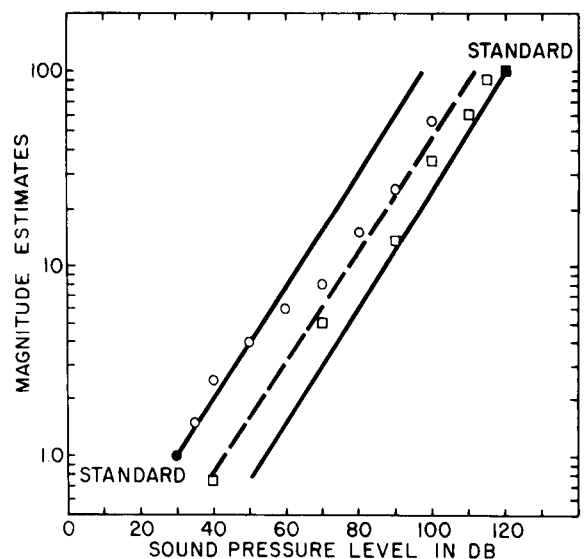


Figure 3. Median magnitude estimates of loudness at 1,000 Hz as a function of SPL. Open circles and squares indicate results obtained with reference standards shown by corresponding filled symbols. The solid lines indicate the expected loudness estimates for the two reference standards. The intermittent line, parallel to the solid lines, joins the lowest loudness estimates associated with the high standard to the highest estimates associated with the low standard. (Data from S. S. Stevens, 1956.)

production scales. Consistent with such a scale was a result of Ward's (1973) experiments in which subjects made multiple loudness estimates of a 1,000-Hz tone relative to a standard at 56 dB SPL called 10. The standard stimulus was identified by the number only at the beginning of the stimulus series, although it was presented repeatedly during the series. At the end of the series, the average number associated with the standard stimulus was 2.08 instead of 10. This number coincides almost exactly with the intermittent line of Figure 3.

It would be naive to conclude that all the biases encountered in ratio scaling procedures can be accounted for by people's tendency toward an absolute scale. Perturbations in the vicinity of the modulus imposed by the experimenter (e.g., Hellman & Zwislöcki, 1961; Stevens, 1956), the regression effect (Stevens & Greenbaum, 1966), the effect of experience (Hellman & Zwislöcki, 1963), and the intramodal range effect (R. Teghtsoonian, 1973) are examples to the contrary. However, these effects cannot account for the partial convergence of the data of Figures 2 and 3 on the scale of Figure 1, or for the predictability of equal sensation magnitudes from similar scales.

To investigate further the existence and robustness of absolute psychophysical scales, we tested in the current study the following propositions, using initially naive subjects. For a given stimulus continuum, (1) different groups of subjects produce approximately the same absolute magnitude-estimation or magnitude-production scale; (2) a scale determined by means of absolute magnitude production on one group of subjects coincides approximately with a corresponding scale determined by means of absolute magnitude estimation on a different group; (3) the scale does not depend on the subjects' experience in psychological scaling without designated moduli; (4) the scale does not depend on the intensity of the first stimulus presented in absolute magnitude estimation or on the first number given in absolute magnitude production; and (5) the scale does not depend on the extent or position of the intensity or number ranges used or available within the confines of these ranges.

The first proposition concerns the generality of the scale; the second, its symmetry; the third implies a natural scale and considers its stability; the fourth differentiates between ratio and absolute scales; the fifth does the same and, in addition, implies that the scale is not an accidental result of a preferred range of numbers.

## METHODS

### General

With the exception of one experimental series in which subjective line lengths were estimated, all the experiments were performed with 20-msec bursts of a 1,000-Hz tone presented mon-

aurally by means of special insert phones. The phones minimize interaural sound transmission and provide high attenuation of ambient sound, including low frequencies. The tone bursts had approximately Gaussian envelopes and were repeated once per second. The short tone bursts were originally selected for experiments in which comparisons with short light flashes were made, and their duration is not relevant to the purpose of the study described here. Their drawback is that they preclude direct comparisons with the older studies discussed above in which much longer tone bursts were used. However, this study is self-contained.

During all the experiments, except those with line lengths, the subjects sat in a soundproofed booth and could neither hear nor see what the experimenter was doing. Every experimental session consisted of three sequences, called runs, during every one of which a full complement of stimuli of the session was presented. As a consequence, every stimulus was judged three times.

In AME experiments, the tone bursts were presented at various SLs in a random order different for every subject, run, and session. The subjects could listen to them as long as they wished before assigning a number. However, they were urged to be spontaneous about their decisions and to make them as quickly as possible. After a number had been assigned, the experimenter interrupted the burst sequence, noted the number, and changed the sound intensity. Thereupon, the bursts were restarted after an appropriate verbal warning to the subject. Prior to some of the experiments, the subjects were introduced to the concept of sensation magnitudes with the help of line lengths drawn on a blackboard. Short and long lines were drawn, and it was explained that there can always be a shorter line than the shortest one drawn and a longer one than the longest drawn. It was added that similar relationships prevail for numbers, in other words, that it is not possible to run out of them. Unless otherwise stated, the subjects were not asked to assign numbers to line lengths, nor did the experimenter name any numbers in his explanation. The use of the lines without assigning numbers to them did not seem to affect the results. After the introduction to the concept of sensation magnitude, the subjects were instructed that they would hear repeated tone bursts and that their task was to assign numbers to the loudness magnitudes of these tone bursts so that the *subjective magnitudes* of the numbers matched the *subjective magnitudes* of loudness. They were instructed to concentrate on each presented burst sequence individually and not to be concerned with the numbers they associated with preceding sequences. They could use any positive numbers that appeared appropriate to them—whole numbers, fractions, decimals. They should not think of any rules they might have learned about numbers, but should be as spontaneous and quick in their responses as possible.

The AMP procedure differed from the AME procedure in that the experimenter gave the numbers and the subjects had to adjust the tone intensity to match the subjective magnitude of loudness to the subjective magnitude evoked by a given number. They made the adjustments by means of a smooth, unmarked knob that did not provide any visual or mechanical cues except for the stops at the two ends of the attenuator range. The stop at the low end of the intensity range was not disturbing, since it was placed well beyond the threshold of audibility. However, the stop at the upper end corresponded to an SPL of about 90 dB and was occasionally reached by the subjects during the adjustment procedure. To minimize the effect of this boundary, we selected the highest number so that, on the basis of AME experiments, it corresponded to a substantially smaller SPL, and we instructed the subjects to tell us if they felt that a given number required a loudness greater than the maximum loudness to which they could turn the knob. Two subjects found that this was the case. Our results suggest that our efforts at overcoming a possible unwanted effect of the upper intensity boundary were reasonably, although perhaps not entirely, successful. To further ensure that subjects could not use the position of the attenuator under their control as a cue, a second attenuator allowed the experimenter to vary the relationship between the position of the subjects' attenuator and sound intensity.

*In every experimental series, the subjects used (different ones in each case) had had no previous experience with magnitude estimation or production procedures, and their experience, if any, with any kind of numerical rating was limited.*

### Specific Experiments

In total, six separate experimental series were conducted in which 78 subjects participated. In the first four and the sixth, AME was used; in the fifth, AMP.

In the first series with 10 subjects, tone bursts at a sound frequency of 4 kHz were scaled jointly with tone bursts at 1 kHz. The experiments consisted of three sessions of six runs each, alternating between the 1- and 4-kHz runs and presented in an order that differed from session to session and subject to subject.

In the second series, performed on eight subjects, the loudness of 1-kHz tone bursts of 20-msec duration was estimated in the first and third sessions and that of 200-msec bursts in the second session.

The third experimental series comprised two sessions with 18 subjects divided into three subgroups of 6 subjects each. The first subgroup received the first stimulus at 6 dB SL in the first session and at 36 dB in the second. The second subgroup started with 36 dB SL in the first session and 78 dB in the second; the third group started with 78 dB in the first session and 36 dB SL in the second. The stimuli of 6 and 78 dB SL were, respectively, the lowest and the highest used in any of the experimental series. In the remaining stimulus presentations, the sound intensities were randomized, as in the other experimental series.

The experiments of the fourth series concerned possible effects of range location. The 12 participating subjects were subdivided into two equal subgroups. One was given a relatively low range of SLs, 6 to 54 dB, in the first session and a relatively high range, 30 to 78 dB, in the second, third, fourth, and fifth sessions. These sessions were followed by three sessions with the low range. For the second group, the order of the ranges was reversed—first the high range, then the low range in four sessions, then the high range in three.

The fifth, and final, experimental series with tone bursts was devoted to magnitude production. The numbers used, derived from magnitude estimation experiments with the other groups of subjects, were .1, .2, .5, 1, 2, 5, and 10. A group of 12 subjects who did not take part in any of the magnitude-estimation experiments participated in two sessions. The subjects were subdivided into two equal subgroups, one receiving .1 as the first number in the first session and 10 in the second session, the second starting with 10 in the first session and .1 in the second session. Otherwise, the numbers were randomized among runs, sessions, and subjects.

The last experimental series concerned lengths of lines. Seven black lines of various lengths drawn on a light background (Life Science Associates, 1970) were projected on a reflecting screen, one at a time in random order, except that neither the shortest nor the longest line ever appeared first. The line lengths on the screen ranged from about 1 to 100 cm. They were viewed in a lighted room at a distance of 3 m by 6 adults and at a distance of 3.7 m by 12 children between the ages of 5 and 6 years. The small difference in distance is unimportant because of size constancy (e.g., M. Teghtsoonian & Beckwith, 1976). The adult subjects were undergraduate students; the children constituted the complete first grade of a private school. Neither group had had any experience in magnitude estimation or production. The children were used because, in contradistinction to the adult subjects, they had a limited range of numbers at their disposal. They knew no fractions, decimals, or numbers greater than 99. Some had heard of the number 100 and occasional higher numbers. The children were subdivided into two equal subgroups, one receiving the second shortest line as the first stimulus, the other, the second longest. Otherwise, the lines were presented in random order. The adults made only one estimate per line since they remembered the numbers too well for a second estimate to be meaningful. The children made two estimates, but only the first is used in this article. The adults were simply instructed to assign numbers to

lines so that the subjective magnitudes of numbers matched the subjective magnitudes of line lengths. They were asked to make intuitive responses and not to think of physical measures of length or rules of arithmetic. The children were introduced to the scaling procedure by the following set of questions to which they could respond "yes" or "no." "Do you know some large numbers? How about some small numbers? Can you think of numbers that are neither small nor large? Can you imagine a line that is long? Can you imagine a short line? How about a line that is neither very long nor very short?" Then they were told: "Now, I am going to show you lines of various lengths on the screen. Do you think you can tell me which numbers fit the lines you will see? O.K. Let's try it. Here is the first line."

## RESULTS

The results are organized around the five propositions stated in the introduction. Those shown in Figure 4 constitute a test of the first proposition that different groups of subjects produce approximately the same magnitude-estimation or production scales. They also provide a partial test of the second proposition that numbers are associated with sensation magnitudes in the same way whether magnitude estimation or production is used. The group geometric means obtained in the three runs of the third session

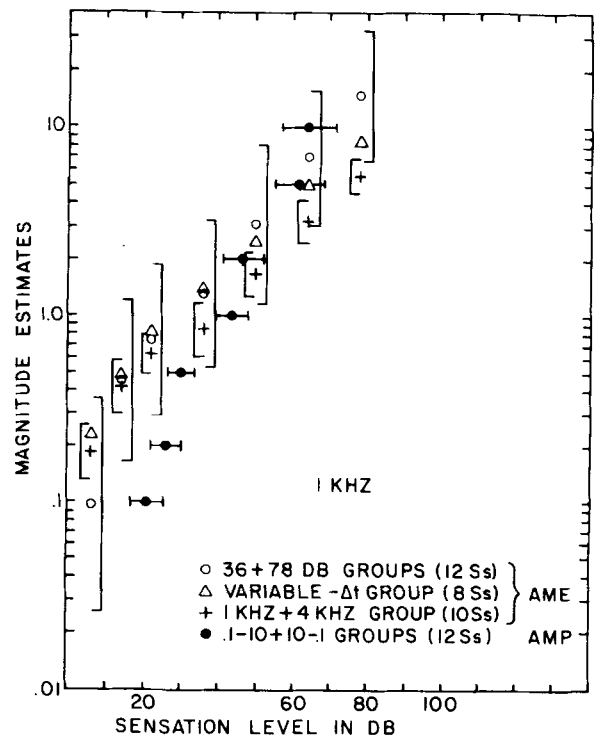


Figure 4. AME results obtained on three groups of moderately experienced subjects in various contexts and AMP results obtained on one group of moderately experienced subjects. The data points show geometric means of unnormalized individual scores; the tall brackets associated with the open circles indicate double standard deviations, and the short brackets associated with the vertical crosses and the horizontal lines through the filled circles, double standard errors of the means.

of the first experimental series are shown by crosses, those of the three runs of the third session of the second experimental series, by triangles. The geometric means of the last two runs of the second session of the third experimental series, obtained on the two subgroups starting with 36 and 78 dB SL, are shown by the unfilled circles. Finally, the filled circles show the means of the last two runs of the second session of the fifth experiment, in which magnitude production was used. The brackets to the left of the crosses indicate their standard errors, those to the right of the circles, the standard deviations of the underlying individual data, and the horizontal lines through the circles, their standard errors. Since the standard deviations shown in Figure 4 are typical of all the experimental series described in this article, they are omitted in the remaining figures concerning loudness. The same is true for standard errors of the means, which are shown in Figure 4 only to indicate the stability of the group means.

It is apparent on sight that the grand means of all four sets of means in Figure 4 are roughly the same, so that all the groups used approximately the same absolute units, irrespective of contextual differences and of whether magnitude estimation or production was required. This satisfies our first and second propositions. On the other hand, the differences seem to have affected the slopes of the implied curves. The flattest slope was obtained when 1- and 4-kHz tones had to be scaled in the same session, in agreement with earlier observations (e.g., Indow & Stevens, 1966), that the slope decreases as the complexity of the task increases. Magnitude production yielded the steepest slope, in agreement with the well-known regression effect (Stevens & Greenbaum, 1966).

The data of Figure 5 address the third proposition, that there should be no effect of experience in magnitude estimation and production without designated standards. The data marked by circles and triangles were obtained in the third experimental series. The geometric means for the entire group are shown by the triangles for the first run of the first session and by the unfilled circles for the second and third runs combined; the filled circles refer to the last two runs of the second session. Thus, the triangles refer to the most inexperienced subjects and the filled circles to the most experienced ones. The effect of experience appears to be both appreciable and systematic. It decreases the assigned numbers at all SLs, particularly at the low ones, and takes place quite rapidly, since the difference between the last two runs of the first and second sessions is small. These findings agree with earlier observations (e.g., Hellman & Zwislöcki, 1963).

The crosses and asterisks in Figure 5 refer to the data obtained with magnitude production in the fifth experimental series. The vertical crosses show the

group means of SLs obtained in the first run of the first session; the slanted crosses, the means of the second and third runs of the same session; and the asterisks, the means of the second and third runs of the second session. Here, the effect of experience is very small, if any. The systematic change apparent in association with the number 10 may be due in part to the finite range of the attenuator, the effect of which the subjects gradually learned to overcome.

Thus, in agreement with the third proposition, experience with magnitude production without designated standards does not affect the scale, but this is not true for magnitude estimation. When subjects have to find sensation magnitudes that match the subjective magnitudes of given numbers, they seem to achieve satisfactory results right away. However, when the numbers are under their control, they first choose ones that later appear too large to them. With experience in magnitude estimation, the numbers converge on those associated with the same SLs in magnitude production. The convergence demonstrates that magnitude production is the more stable procedure and suggests that the association of numbers with sensation magnitudes is asymmetrical in inexperienced subjects. This may be so because numbers, but not sensations, are learned and, formally, numbers have no absolute values. Pragmatically, one

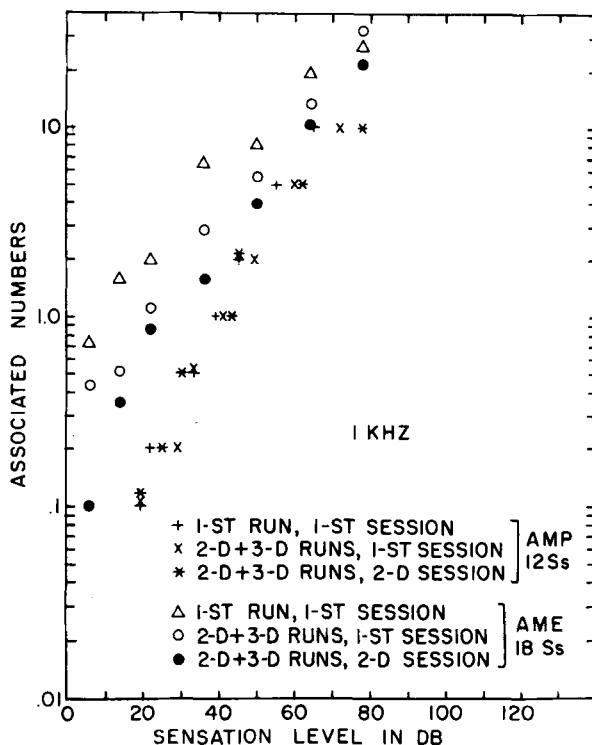


Figure 5. Geometric means of AME and AMP results obtained on two groups of subjects, respectively, at three levels of experience. The triangles and vertical crosses correspond to the lowest level, and the filled circles and asterisks, to the highest.

would like to suggest magnitude production as the method of choice; unfortunately, it is often difficult to implement. Instead, it is shown below how the effect of experience in magnitude estimation can be minimized.

The test of the fourth proposition, that the initial sensation magnitude or number should not affect the scale, is illustrated in Figures 6 and 7. The third experimental series on magnitude estimation is represented in Figure 6, which shows the geometric means from the first run of the first session. Filled circles indicate the subgroup whose first stimulus was 6 dB SL; the unfilled circles, 36 dB; and the crosses, 78 dB. Arrows point to the geometric means of the very first estimates ever made by the subjects. The data points show that, within each subgroup, the mean numerical estimates of the initial stimuli are consistent with the mean estimates of the subsequent stimuli and, furthermore, that the mean estimates of the initial stimuli are consistent with the means of the subsequent estimates of the entire group. The initial stimuli had a small, but systematic, effect on the loudness function, however. Subjects who started with 6 dB SL tended on the average to assign higher numbers than subjects who started with 36 dB SL, and those starting with 78 dB SL tended to assign the lowest numbers. Also, the first subgroup produced a steeper loudness curve than the remaining two. The biases are consistent with the regression effect, which has been blamed on people's reluctance to use extreme values. Accordingly, the numbers chosen by the subgroup receiving their initial stimulus at an extremely low SL are higher than the num-

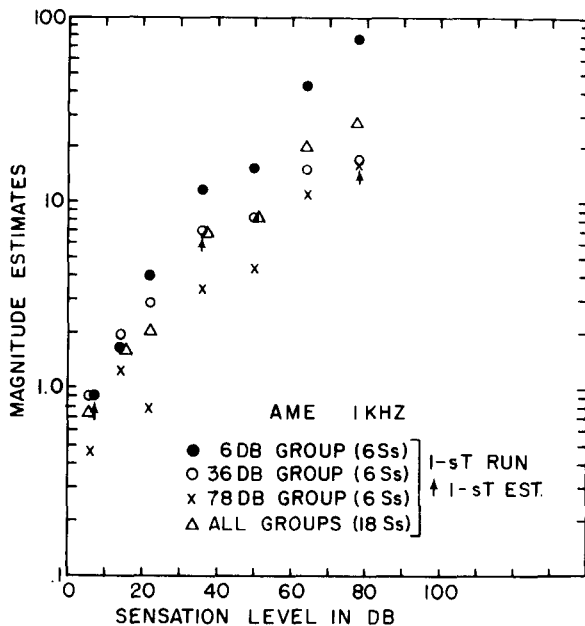


Figure 6. Geometric means of AME scores of inexperienced subjects in three subgroups who received the initial stimulus at three different sensation levels. The arrows point to the means of the very first estimates.

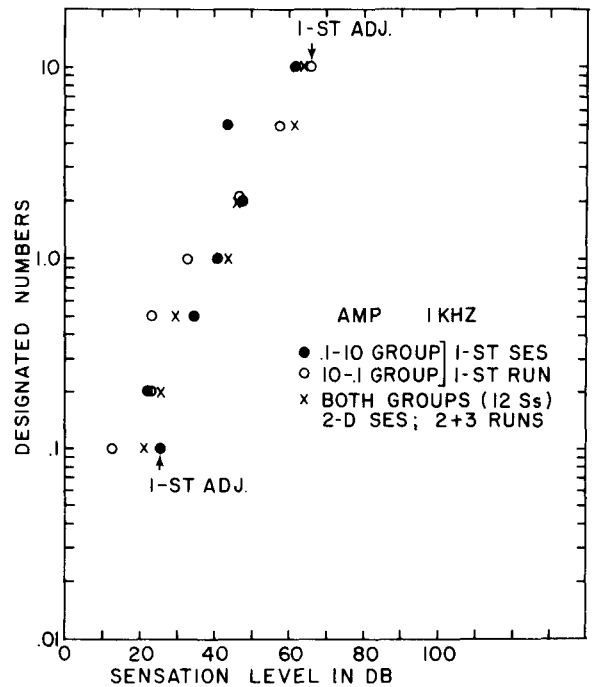


Figure 7. Mean SL results obtained with AMP on two groups of initially inexperienced subjects, one of which received .1 as the first number, and the other, 10. The circles refer to the first run of the first session, and the arrows point to the very first adjustments. The crosses are the means of the results obtained in the last two runs of a second session on both groups combined. In that session, the first group started with the number 10 and the second with the number .1.

bers chosen by either of the remaining two subgroups. The pattern of data points in Figure 6, as well as in Figure 5, also suggests that the subjects were reluctant at first to use numbers smaller than 1. Nevertheless, it should be clear that the first estimates were not made at random (ratio scale) but roughly followed the loudness function for inexperienced subjects.

The effect of the initial number given the subjects in magnitude production is shown in Figure 7. Filled circles indicate the subgroup that received the initial number .1; unfilled circles, the subgroup that received the initial number 10. Both sets show means of SL adjustments in the first run of the first session. For comparison, crosses indicate the means from the last two runs of the second session. In this session, the group that had received the low numbers as the initial number in the first session began with the high number, and the remaining group began with the low number. The three sets of results show that neither initial number nor experience produced any systematic effect on the loudness function. This is consistent with the data of Figure 5 in demonstrating the stability of the scales obtained by means of magnitude production.

The results shown in Figures 6 and 7 taken together satisfy roughly the fourth proposition, that the scale should be independent of the initial stim-

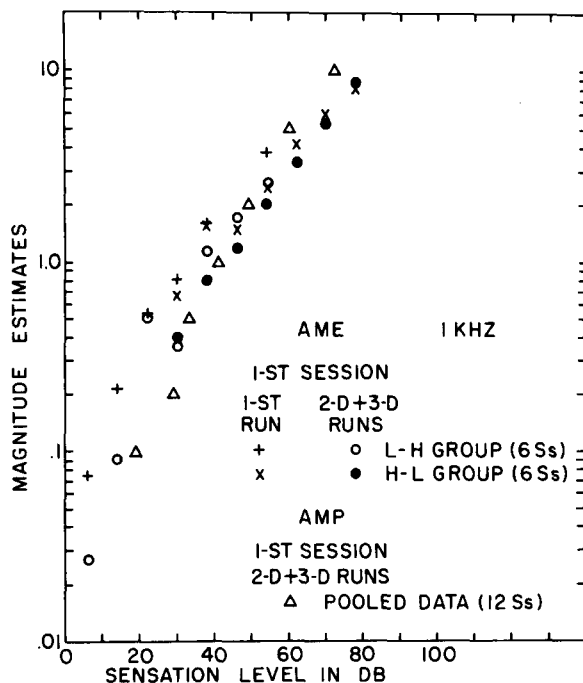


Figure 8. Geometric means of AME results obtained on two groups of subjects receiving different ranges of SLs, 6 to 54 dB and 30 to 78 dB. The crosses refer to the first run of the first session, the circles to the last two runs of the same session. For comparison, the triangles show the mean AMP results obtained on a third group of subjects in the last two runs of their first session.

ulus or number given the experimental subjects. The bias evident in Figure 6 is probably related to the bias discussed in connection with Figure 5 and seems to be due to the familiar regression effect.

The fourth experimental series was intended as a test of the fifth proposition, that an absolute sensation scale does not depend on the location of the range of the stimuli used. To minimize the initial effect of experience on magnitude estimation, the subjects were asked to assign numbers to some lines drawn on the blackboard. As a first line, a medium-length line was drawn. Subsequently, some very short lines were drawn to induce the subjects to use fractions or decimals, since previously discussed observations had suggested that many subjects were reluctant at first to use numbers smaller than 1. One of the two participating groups received in the first session a range of intensities between 6 and 54 dB SL, and the other, a range between 30 and 78 dB SL. In the second session, the ranges were interchanged and were maintained the same in three subsequent sessions. For an additional sequence of three sessions, they were interchanged again. Accordingly, the first group (L-H) went through the sequence L-H-HHH-LLL; for the second group (H-L), the sequence was H-L-LLL-HHH.

The geometric means of the magnitude estimates obtained in the first session are shown in Figure 8.

The crosses refer to the first run and the circles to the last two runs of the session. Note that the difference between initial and subsequent estimates is much smaller than in Figure 5. In addition, the geometric means of the second and third runs agree well with the corresponding means of magnitude production (triangles) transferred from Figure 5 for comparison. We ascribe this stabilization of magnitude-estimation results to the informal scaling of a few line lengths prior to scaling loudness.

The data shown in Figure 8 reveal no range-location effect, in complete agreement with the fifth proposition. In the overlapping range of intensities, both groups associated practically the same numbers with the same SLs, and, outside this range, the data fall on nearly the same implied curve. Although somewhat higher numbers were used in the first run than in the subsequent two, the difference was roughly the same for both groups. It is clear that both groups applied from the start the same average units, independent of the range location.

The effect of changing the range location from low to high, or vice versa, is shown in Figure 9. The data points refer to geometric means of the last two runs

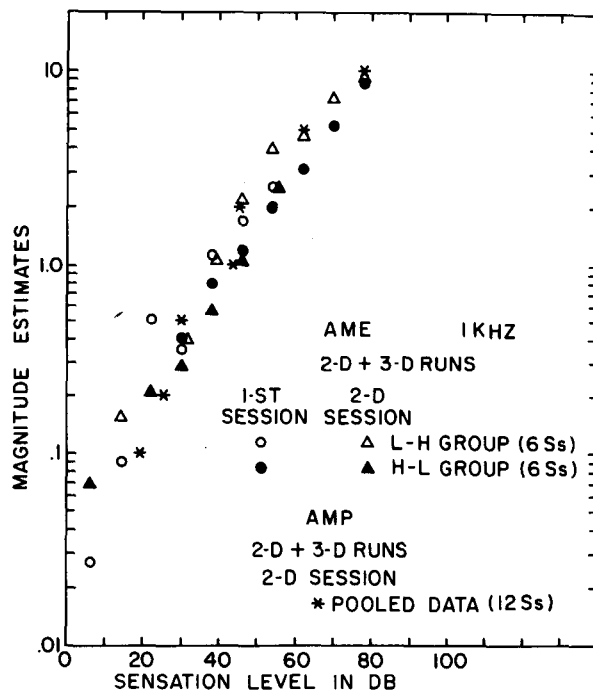


Figure 9. Comparison of geometric mean results obtained with AME on two groups of subjects receiving different ranges of stimulus SLs. The circles refer to the last two runs of the first session, and the triangles, to the last two runs of the second session. The open circles belong to the group that received the low range in the first session and the high range in the second, the closed circles to the group for which the ranges were presented in a reversed order. For comparison, the asterisks show the mean AMP results obtained on a third group of subjects in the last two runs of their second session.

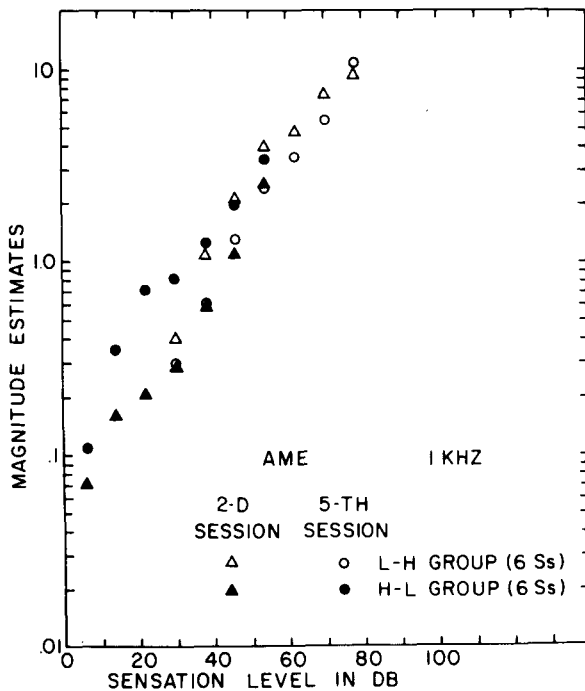


Figure 10. Comparison of geometric mean results obtained with AME on the same two groups of subjects as in Figures 8 and 9. The triangles refer to all three runs of the second session, and the circles to all three runs of the fifth session. The unfilled symbols belong to the group receiving the high range in all the sessions, from second to fifth, and the filled symbols to the group receiving the low range in the same sessions.

of the first and second sessions, respectively. The unfilled symbols belong to the L-H group and the filled ones to the H-L group. No effect of range change is detectable. (The data of the first run of the second session were omitted because, as in the first session, slightly higher numbers were used in that run. However, the difference was not contingent upon the intensity range.) For comparison purposes, asterisks in Figure 9 indicate means of sensation levels derived from the second and third runs of the second session of the AMP experiment. Their almost exact agreement with the corresponding AME data demonstrates that different groups of moderately experienced subjects used the same absolute units in both procedures. This satisfies our second proposition.

It may be worth noting that Group L-H tended to produce somewhat higher numbers than Group H-L, whether it was given the low or the high range. The difference is well within the intergroup variability pointed out in connection with Figure 4.

We further investigated the possibility that repeated presentation of the same stimulus range over several sessions affects AME results. Geometric means of magnitude estimates obtained in all three runs of the second session are compared with those of the fifth session in Figure 10. A small, but systematic, effect is evident. The group receiving the

high range repeatedly used somewhat smaller numbers in the fifth session than in the second one (unfilled symbols), and the reverse was true for the group receiving the low range (filled symbols). These changes are consistent with a growing tendency toward the use of the same number range independent of the stimulus range. We shall limit our speculations about the reasons for this and simply state that the subjects seemed bored by the repetition of an identical task and that excessive repetition of AME, using the same stimulus range, does not improve performance.

When the ranges were interchanged again between the two subgroups in the sixth session, the effect of prolonged experience with a different range seems to have disappeared, as is evident in Figure 11. The unfilled symbols indicate the geometric means of the last two runs of the first session and of all three runs of the sixth session by the group receiving the low range in these sessions. The filled symbols show corresponding results for the group receiving the high range. The first run of the sixth session is included in the geometric means of that session, since it did not differ systematically from the remaining two runs. Note that the subjects receiving the low range assigned on the average almost identical numbers in the first and sixth sessions, in spite of the interposition of four sessions in which they were ex-

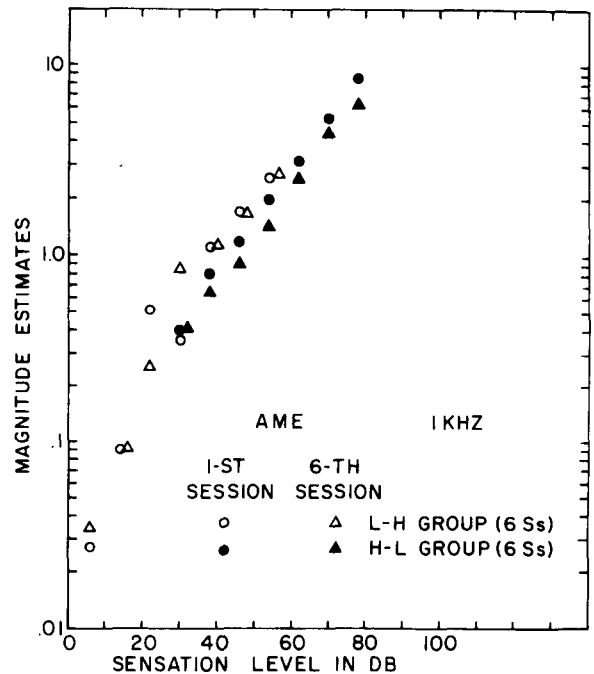


Figure 11. Comparison of geometric mean results obtained with AME on the same two groups of subjects as in Figures 8 through 10. The circles refer to the last two runs of the first session, and the triangles to all three runs of the sixth session. The unfilled symbols belong to the group receiving the low range in both sessions, the filled symbols to that receiving the high range in the same sessions.



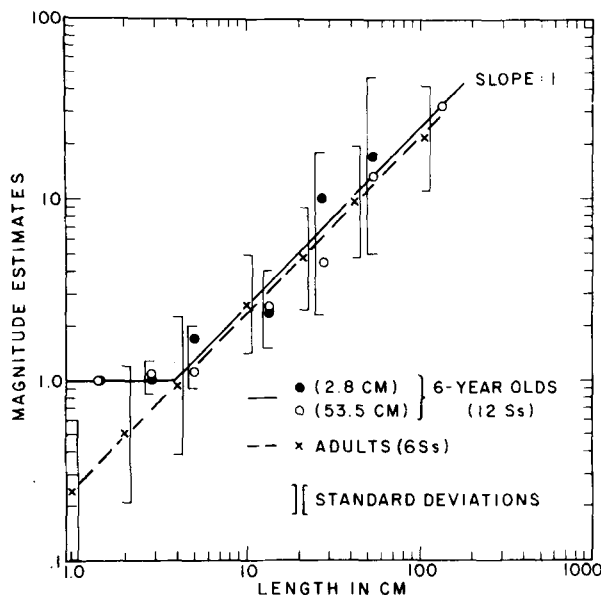


Figure 12. Geometric means of AME of line lengths of adults (crosses) and 5- to 6-year-old children (circles). The filled circles show children receiving 2.8 cm as an initial stimulus, and the unfilled circles, 53.5 cm. Brackets indicate standard deviations. Means are based on estimates from the first run, first session.

posed to the high range. The other group assigned slightly lower numbers in the sixth session than in the first, which may be blamed either on increased experience per se or on the interposed four sessions with the lower SL range. This group also showed a relatively stronger effect of range repetition (filled symbols in Figure 10). However, the systematic effect evident in Figure 11 is small. On the basis of the results shown in Figures 8 to 11, we conclude that the range location of stimulus intensities has little, if any, effect on AME results. In other words, the subjects use approximately the same units independent of the location of the stimulus range.

The effect of number range was investigated in the last experimental series on line lengths, using as subjects adults whose number range was unlimited and children who knew numbers only in the range between 1 and 99. The geometric means of the data are compared in Figure 12. The filled circles refer to the subgroup of six children who received a 2.8-cm length as their first stimulus; the unfilled circles refer to the remaining six children who started with the 53.5 cm. The crosses refer to the adults. The lines were fitted to the data by eye, and the brackets indicate the standard deviations of log scores relative to the geometric means of the two groups. There is no systematic difference between the scales produced by the adults and by the children within the range of numbers available to the latter. Both scales are consistent with those obtained by M. Teghtsoonian and Beckwith (1976), who used 18-year-old adults and children ranging in age from 8 to 12. The limited

number range at both ends of the children's scale does not seem to have affected either the slope or the position of the scale. Of particular interest is the fact that the starting line length had almost no effect on the children's scale, even though the adults assigned a fractional number to the shorter one of the two initial lines, and the children, not knowing fractions, had to use the number 1. Note the truncation of the children's scale at line lengths to which the adults assigned fractional numbers. The children used for these lines the smallest number they knew, which was 1. Some hesitated between 1 and 0, but ultimately chose the former.

## CONCLUSIONS

By means of six experimental series, we investigated the validity of the hypothesis that, in magnitude-estimation and production procedures without designated moduli, people tend to scale their sensation magnitudes on absolute, rather than ratio, scales. More specifically, we tested the following propositions: (1) Different groups of subjects produce approximately the same mean scale; in other words, they produce approximately the same function relating numbers associated with sensation magnitudes and the corresponding physical stimulus magnitudes. (2) A scale determined by absolute magnitude estimation (AME) on one group of subjects coincides approximately with the scale determined by absolute magnitude production (AMP) on a different group of subjects (when neither group is familiar with the complementary method). (3) The scale does not depend on the subjects' experience with AME or AMP procedures. (4) The scale does not depend on the magnitude of the first stimulus or number given to the subjects. (5) Within the available stimulus or number ranges, the scale does not depend on their location or extent.

Our results show that different groups of subjects do produce nearly equal mean scales when no reference moduli are imposed and, consequently, the subjects are free to choose their own units (Figure 4). However, the units vary from subject to subject over a range of more than one order of magnitude, producing geometric standard deviations of about 7 at medium and high magnitudes and ones even greater at low magnitudes to which small fractional numbers are assigned. This means that, if people do use absolute scales, the scales are private and different from each other. Nevertheless, the units show a sufficient central tendency, so that mean scales produced by small groups of about 10 subjects agree within a ratio of units on the order of 2.

That people tend to use absolute rather than ratio scales is demonstrated by the fact that, on the aver-

age, the units chosen depend little on the first stimulus scaled (Figure 6), on the location of the stimulus range within the variable continuum (Figures 8 and 9), on the extent of the range (compare Figures 1 and 9), and even on the range of available numbers (Figure 12). If people used a constant, preferred range of numbers, which they map onto the stimulus range, whatever that range may be, variation of any one of these parameters would lead to different units. For instance, if people used the first stimulus presented as a unit and assigned to it the number 1, all the initial estimates shown in Figure 6 by arrows would line up on the ordinate of 1. Instead, the subjects in our experiments chose a small mean number when the initial stimulus was weak, a relatively large number when it was strong, and an in-between number when it was of medium strength, even though they had no experience with scaling whatsoever at the time they made their first numerical estimates. Similarly, when two groups of inexperienced subjects received two different stimulus ranges, one extending from 6 to 54 dB SL and the other from 30 to 78 dB SL, they did not use the same range of numbers for both, which would have led to different units, but chose the numbers so as to maintain the same unit (Figure 8). Exchange of the ranges between the two groups left the unit invariant (Figure 9). Perhaps the most striking result pointing to an absolute, rather than a ratio, scale is the effect of a limited number range shown in Figure 12. Although the children involved in the experiment did not know fractions, they did not decrease their units to avoid using them. On the contrary, they used the same average unit as the adults, which forced them to truncate their scale at stimulus values the adults associated with fractions. The unit proved so robust that it remained practically invariant even when the magnitude of the first stimulus was clearly below the point of truncation.

The fact that the children used the same average unit as the adults points to the stability of the unit and to its origin. When children learn numbers, they do not learn them at first as abstractions but by counting objects. Thus, the original coupling of numbers with the impressions of physical quantities occurs through numerosity. This has two consequences, since a numerosity scale is an absolute scale: no fractions occur and the units are fixed.

Scaling subjective line length and other sensation continua requires a transfer from the numerosity scale, which allows an arbitrary definition of scale units. As Figure 12 demonstrates, even young children can make the transfer without difficulty, and the average unit they choose is the same as that chosen by the adults. On the other hand, the inherent arbitrariness of choice appears to be reflected in the great intersubject variability (Figures 4 and 12). As already discussed, it does not lead to a substantial

intrasubject variability, however. On the contrary, the individual units, once formed, tend to remain reasonably stable and, at least on the average, practically invariant from the time a sufficient range of whole numbers is learned.

If the subjective units originate from numerosity, in which fractions do not occur, it is possible to propose a plausible explanation for the increased instability of magnitude-estimation scales at low stimulus levels, which is evident in Figures 4 through 6. The first of these figures shows an increased intersubject variability at the stimulus levels associated with fractional numbers. The second and third figures suggest an initial reluctance on the part of the subjects to use fractional numbers. The instability can be eliminated either by using magnitude production, in which numbers are given by the experimenter (Figure 5), or by allowing the subjects first to scale some medium and very short line lengths to which they assign fractional numbers. Once the subjective "fraction barrier" is broken, the scale stabilizes and becomes practically independent of experience with magnitude estimation without designated moduli (Figures 8 through 11).

After the "fraction barrier" has been broken, the magnitude-estimation scale comes to coincide nearly exactly with the magnitude-production scale (Figure 8), which has proved to be remarkably stable with respect to both experience (Figure 5) and the initial designated number (Figure 7). Under these conditions, the coupling between numbers and sensation magnitudes is symmetrical, that is, independent of whether the numbers or sensation magnitudes are under the subjects' control. The symmetry is preserved even when magnitude estimation is performed by one group of subjects and magnitude production by another, neither group having any knowledge of the complementary procedure.

Our demonstration that subjects, when left unconstrained, tend to scale sensation magnitudes on absolute scales does not preclude the possibility of ratio scaling within limited stimulus ranges. This follows from Figures 2 and 3. Whether subjects accept the units imposed by the experimenter or use their own preferred units depends on instructions and the subjects' capabilities. The instructions with which the results of the present study were obtained asked the subjects to match the subjective magnitudes they associated with number symbols to their sensation magnitudes on a one-to-one basis, disregarding any context that may have been present. In agreement with earlier suggestions (Hellman & Zwislocki, 1961; Stevens, 1959), these instructions imply that association of psychological magnitudes evoked by number symbols with other psychological magnitudes is only a special instance of intermodality matching. In general, such matching occurs in absolute terms.

The following is a specific example of our instructions for AME of line lengths.

In this experiment we would like to find out how large various lengths of lines appear to you. For this purpose, I am going to show you on the screen a series of lines of various lengths, one at a time. Your task will be to assign a number to every line in such a way that your impression of how large the number is matches your impression of how large the length of the line is. You may use any positive numbers that appear appropriate to you—whole numbers, decimals, or fractions. Don't think of physical units of measurement, such as inches or centimeters, and do not worry about running out of numbers—there will always be a smaller number than the smallest you use and a larger one than the largest you use. Do not worry about numbers you assigned to preceding lines. Do you have any questions?

Note that we avoided asking subjects to assign numbers to line lengths or to match the subjective impression of how large a number is to the subjective impression of how long a line is. Matching can occur only within a common domain, and we assume that both a number symbol and a line length evoke an abstract magnitude impression that is common to both. This appears consistent with R. Teghtsoonian's (1973) suggestion that sensation magnitudes have a common range.

In addition to their fundamental significance, the experimental results we have described in this article have practical applications. Because the average absolute scales seem to be influenced little by the range of the magnitudes involved and to exhibit remarkable invariance among experimental groups, they should have widespread validity. For example, loudness of noise scaled at some location in New York City on one group of people should be roughly comparable with the loudness of noise scaled at some location in Los Angeles by a different group of people.

#### REFERENCES

- 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., & 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 at 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.
- INDOW, T., & STEVENS, S. S. Scaling of saturation and hue. *Perception & Psychophysics*, 1966, **1**, 253-271.
- KRANTZ, D. H. A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, 1972, **9**, 168-199.
- LIFE SCIENCE ASSOCIATES. *Magnitude estimation, length judgments*. Box 163, Baldwin, N.Y. 11510: Author, 1970.
- MARKS, L. E. Binaural summation of the loudness of pure tones. *Journal of the Acoustical Society of America*, 1978, **64**, 107-113.
- REYNOLDS, G. S., & STEVENS, S. S. Binaural summation of loudness. *Journal of the Acoustical Society of America*, 1960, **32**, 1337-1344.
- ROWLEY, R. R., & STUDEBAKER, G. A. Monaural loudness-intensity relationships for a 1000 Hz tone. *Journal of the Acoustical Society of America*, 1969, **45**, 1193-1205.
- SCHARF, B., & FISHKEN, D. Binaural summation of loudness: Reconsidered. *Journal of Experimental Psychology*, 1970, **86**, 374-379.
- STEVENS, S. S. Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
- STEVENS, S. S. The measurement of loudness. *Journal of the Acoustical Society of America*, 1955, **27**, 815-820.
- STEVENS, S. S. The direct estimation of sensory magnitudes—loudness. *American Journal of Psychology*, 1956, **69**, 1-25.
- STEVENS, S. S. Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology*, 1959, **57**, 201-209.
- STEVENS, S. S. *Psychophysics* (G. Stevens, Ed.). New York: Wiley, 1975.
- STEVENS, S. S., & GREENBAUM, H. Regression effect in psychophysical judgement. *Perception & Psychophysics*, 1966, **1**, 439-446.
- TEGHTSOONIAN, R. Range effects in psychological scaling and a revision of Stevens' law. *American Journal of Psychology*, 1973, **86**, 3-27.
- TEGHTSOONIAN, M., & BECKWITH, J. B. Children's size judgments when size and distance vary: Is there a developmental trend to overconstancy? *Journal of Experimental Child Psychology*, 1976, **22**, 23-39.
- VERRILLO, R. T., FRAIOLI, A. J., & SMITH, R. L. Sensation magnitude of vibrotactile stimuli. *Perception & Psychophysics*, 1969, **6**(6A), 366-372.
- WARD, L. M. Repeated magnitude estimations with a variable standard: Sequential effects and other properties. *Perception & Psychophysics*, 1973, **13**, 193-200.

(Received for publication April 9, 1979;  
revision accepted April 30, 1980.)