Regression effect in psychophysical judgment[1]

S. S. STEVENS AND HILDA B. GREENBAUM[2]
HARVARD UNIVERSITY

Psychophysical judgment, like all other kinds of judgment, involves a matching or equating of two different domains. When the judgment involves the matching of values on two perceptual continua, the observer tends, on the average, to constrict the range of his adjustments on whichever variable is placed under his control. When the observer adjusts each variable in turn, two different regression lines are produced. This regression effect presumably occurs whenever the results of the matching judgments yield less than a perfect correlation. Illustrative examples are given for the continua, loudness, vibration, brightness, and duration.

Psychophysical judgment is typically an operation in which a person matches an aspect of one domain to an aspect of another. Even the act of placing stimuli into categories, such as seen or not seen, can be regarded as an operation in which an element of one domain is coupled, conjoined, or equated to an element of another domain. This universal feature, which makes matching the paradigm of the judging process, has been dealt with more fully elsewhere (Stevens, in press). The purpose here is to examine a feature of judgment that exhibits itself most clearly when the two domains to be conjoined are perceptual attributes produced by stimuli that may be regarded as continuously variable. The observer's task is to vary one or the other stimulus in order to equate in some respect or other the two perceptual attributes.

It has long been observed that, when a person varies one stimulus to match a variable criterion of some sort, he tends to shorten or constrict the range of his adjustments. More than half a century ago, Hollingworth (1910) wrote about the "central tendency of judgment," the tendency for the observer to regress toward a central value and thereby shorten the range of the adjusted variable. He said, "Judgments of time, weight, force, brightness, extent of movement, length, area, size of angles, have all shown the same tendency to gravitate toward a mean magnitude . . ."

A similar ubiquitous regression tendency makes itself felt in the matching experiments by which the psychophysical power law may be demonstrated. Depending on which variable the observer controls, the matching procedures yield two different regression lines (in log-log coordinates) and hence two different exponents. The difference between the values of the two exponents may be very small if the matching judgments are relatively precise and noise-free, or it may be large if the judgments are subject to perturbations and asymmetrical constraints. Matching functions obtained in a wide variety of experiments are illustrated below.

LOUDNESS MATCHING

Judgments made within the auditory sense modality often call for the matching of two stimuli that differ in frequency or in shape of spectrum. Figure 1 shows the results of matching a tone of 1000 Hz and a band of noise 500-2000 Hz (Stevens, 1956). Each stimulus served as both standard and variable in a counterbalanced design. Since the coordinates of Fig. 1 are logarithmic, the slope of the matching function determines the exponent. The exponent clearly depends on whether the tone or the noise was adjusted. Since the noise band was limited to two octaves, the matching could be made with fair precision, and the regression effect is relatively slight. Nevertheless, the difference in the slope of the two lines is evident. The average quartile deviation was 4.4 dB when the tone was adjusted and 3.3 dB when the noise was adjusted.

A factor that may contribute to the difference in slope (exponent) when a wide range of sound levels is used may be the tendency of observers to avoid sounds that seem disagreeable or painfully loud. At the other extreme, faint sounds that a person must strain to hear are also relatively unpleasant as stimuli. The "most confortable listening level" for the typical listener lies

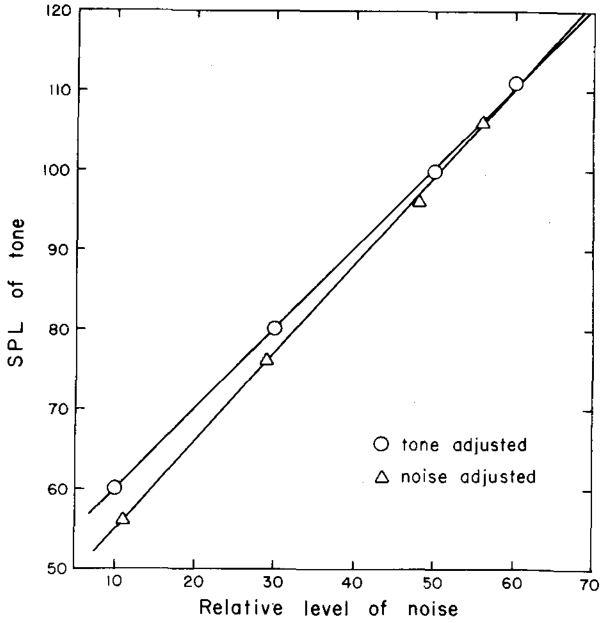

Fig. 1. Loudness balances between a tone of 1000 Hz and a band of noise 500-2000 Hz. Each of 19 observers adjusted the tone to match the noise and the noise to match the tone. The points are the geometric means (decibel averages) of 38 adjustments.

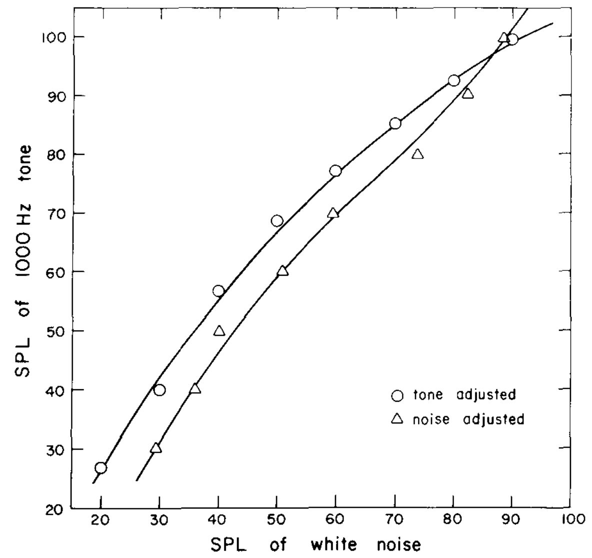

Fig. 2. Loudness balances between a tone of 1000 Hz and a wide-band white noise. Twenty observers matched tone to noise and noise to tone.

in the vicinity of 70 dB SPL, and it is possible that departures from the most comfortable level are resisted by the observer when he has control of the stimulus. A hypothetical comfort factor can account for no more than a part of the regression effect, however, for many judgmental matching tasks do not create discomfort.

Much more formidable than matching a tone to a relatively restricted band of noise is the task of matching the tone to a wide band of so-called white noise. The divergencies among the results from various laboratories attest the uncertainty of the judgment when a tone and wide-band noise are used (cf. Stevens, 1955). In making the matches, most investigators have allowed the observers to vary only the tone, but Zwicker (1958) allowed them to vary both the tone and the noise, with the results shown in Fig. 2. We note that the relation between the loudness of the tone and the wide-band noise is clearly curvilinear in the decibel coordinates, but, despite the curvature, the function obtained when the noise was adjusted is generally the steeper of the two. In other words, the curvature does not obscure the regression effect.

It is clear in Figs. 1 and 2 and in the results of other experiments that the two regression functions do not always cross at their center points. The regression effect is not the only source of systematic error in matching experiments.

MONAURAL-BINAURAL MATCHING

Even when the same white-noise stimulus is used, but a different combination of receptors is stimulated, the regression effect may assert itself. The circles and squares in Fig. 3 show results obtained in 1954-55 by E. Tulving in this laboratory. Seventeen observers adjusted the level of a white noise heard binaurally to match the loudness of the same noise heard monaurally. Some time later, 15 observers reversed the procedure and adjusted a monaural noise to match a binaural noise. In both experiments, the monaural noise was given half the time in the right and half the time in the left ear. The second experiment was actually undertaken to resolve a discussion between Tulving and one of the present authors concerning the validity of a unidirectional loudness-balancing procedure. It appears from Fig. 3 that systematic constraints may act to bias a loudness balance if only one stimulus parameter is varied by the observer.

BRIGHTNESS MATCHING

In an experiment on the matching of the brightnesses of flashes that differed only in duration, a small but systematic regression effect was found (Aiba & Stevens, 1964). Wide ranges of duration and luminance were explored in order to map out the form of the Broca-Sulzer effect. The observers varied the level of a 1-sec. flash to match the brightness of shorter flashes, and they also varied the shorter flashes to match the 1-sec. flash. In log-log coordinates the slopes of the regression lines differed by about 4 percent.

CROSS-MODALITY MATCHING

Incidental to a series of experiments designed to measure the difference between monaural and binaural loudness, Reynolds and Stevens (1960) obtained several

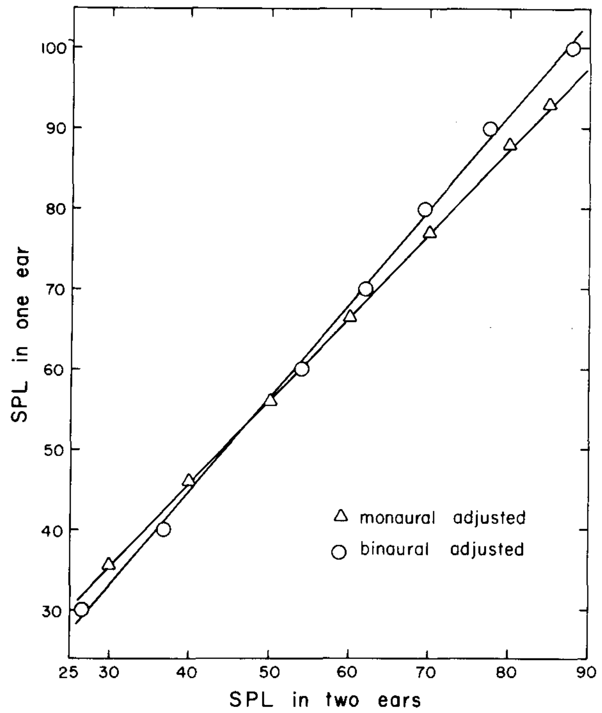

Fig. 3. Monaural-binaural loudness matching. The stimulus was a white noise heard alternately in one ear and in both ears. Seventeen observers adjusted the binaural sound to match the monaural sound. Fifteen observers did the reverse.

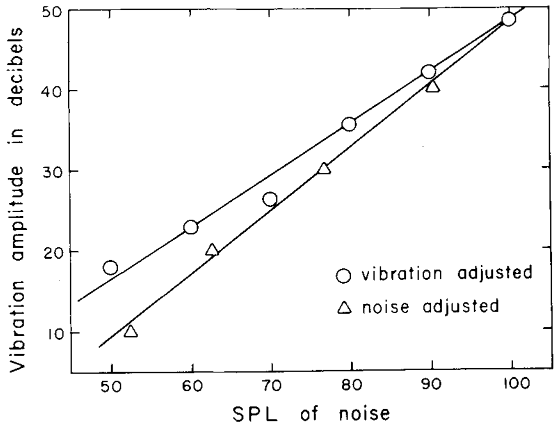

Fig. 4. Cross-modality matching functions between vibration applied to the fingertip and the loudness of a noise 100-500 Hz heard binaurally. Points are based on the decibel averages for 10 observers.

examples of the regression effect. Their main finding was that, although listening with two ears gives an impression of loudness that grows as a power function of sound pressure with an exponent of 0.6, listening with one ear gives an exponent of 0.54. The intercepts of the two loudness functions are such that at 90 dB SPL the loudness in two ears is double the loudness in one ear.

A replotting of some of the functions obtained by Reynolds and Stevens illustrates how the regression effect operates in cross-modality matching. Figure 4 shows the two functions obtained from matches between the loudness of a band of noise 100-500 Hz and a 60-Hz vibration applied to the tip of the middle finger. The slope (exponent) depends on whether the noise or the vibration was adjusted by the 10 observers. One of the constraints in this particular matching experiment results from the relatively narrow range of vibration amplitudes that can be effectively produced. The usable range is limited to about 40 dB. With noise, on the other hand, the usable range can exceed 100 dB if required. For Fig. 4 the noise was binaural. When the noise was monaural a similar pair of regression functions was obtained, but both slopes were less steep.

The nature of the continuum produces less constraint if, instead of matching vibration to loudness, the observers are required to match numbers to loudness. In so-called magnitude estimation the observer matches numbers to a sensory impression under instructions to preserve a proportionality between the numbers and the perceptual magnitude. In magnitude production the procedure is reversed: numbers are given in irregular order and the observer adjusts the stimulus to produce a match. The outcome of the two procedures, with the noise heard binaurally, is shown in Fig. 5. Magnitude estimation gives a flatter function and hence a lower exponent; magnitude production gives a steeper function, with a higher exponent. When the listening was monaural the two matching functions resembled Fig. 5 as regards regression, but both slopes were less steep. It can be seen, therefore, that the procedure of matching numbers to stimuli or stimuli to numbers can be regarded as an instance of the general method of cross-modality matching. There is nothing especially different about number as a perceptual continuum. When a matching experiment involves the number continuum, we find the same regression effect that characterizes other perceptual continua.

The regression effect in number matching also shows itself in experiments designed to scale the inverse attributes. Thus observers may be instructed to match numbers to the softness of tones, instead of to the loudness. The result usually approximates a power function with an exponent that is the negative of the exponent for loudness. The measured value of the exponent differs, depending on whether the observers match numbers to softness or softness to numbers. An example from a study by Stevens and Guirao (1962) is shown in Fig. 6. In addition to the obvious regression effect, there is a noticeable tendency for the data in Fig. 6 to be slightly concave downward. This downward concavity has been a feature, more or less prominent, in the measurements of inverse attributes on some dozen different perceptual continua.

BRIGHTNESS VERSUS LOUDNESS

An experiment designed to match perceptual intensities in the two receptor systems, vision and audition, is especially interesting. J. C. Stevens and Marks (1965) asked 10 observers to adjust the level of a band of noise 75-4800 Hz to make its loudness seem as great as the brightness of a given luminous target subtending 4°. The luminance was varied irregularly over eight values between 50 and 100 dB re 10^{-10} L. In another session the level of the sound was set by the experi-

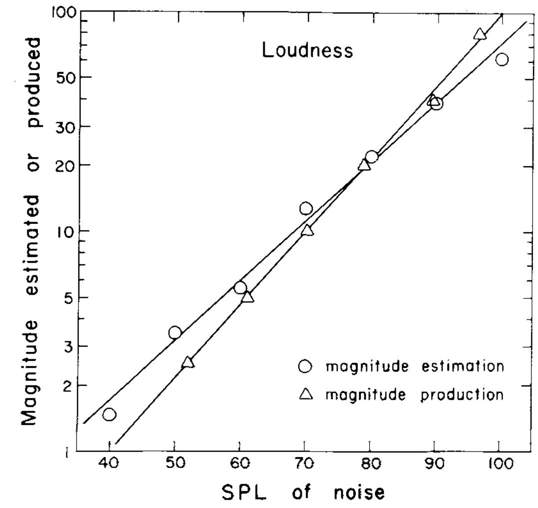

Fig. 5. The matching of numbers to loudness (magnitude estimation) and loudness to numbers (magnitude production). The stimulus was a band of noise 250-2000 Hz. Ten observers made matches.

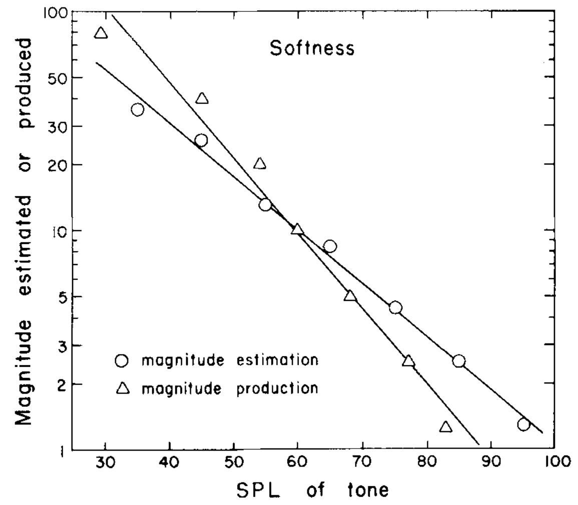

Fig. 6. Matching functions between number and an inverse attribute, the softness of a tone of 1000 Hz. Each point is a geometric mean of 20 judgments, two by each of 10 observers.

menter, and the observer adjusted the level of the luminance to make the brightness match the loudness. In order to keep the eyes in a state of adaptation well below the stimulus levels, the observers were dark adapted with red goggles for at least 10 min. In addition, the visual stimuli were presented for 0.45 sec., and about 8 sec. elapsed between presentations.

The results plotted in Fig. 7 show the expected regression effect. In a second experiment on brightness-loudness matching, the size of the regression effect was somewhat diminished, perhaps because the observers were able, by pressing keys, to see the light and hear the noise whenever they pleased. In the first experiment, the forced wait of more than 8 sec. between stimuli may have strained patience and memory. The usual effect of adding to the difficulty of a matching judgment is to increase the angle between the regression lines.

In most psychophysical experiments only one of the two regression lines is obtained, and under those circumstances the effect of adding difficulty, distraction, or other "noisy" impediments usually shows up in the results as a decrease in the slope (exponent) of the psychophysical power function.

DURATION

The following series of experiments was designed to explore judgments of duration by several different procedures, each of which may be expected to exhibit the regression effect.

Time intervals were marked for the observer by a 15-W red light and by a white noise from a loudspeaker. Appropriate switches and timers allowed the experimenter to present a known duration of one stimulus—light or sound. The observer could try to match the apparent duration by controlling the actual duration of the other stimulus. Durations could be measured to 0.01 sec. The nature of the observer's task is outlined by the instructions:

> You will hear a white noise of constant intensity for varying durations. Shortly thereafter, a red light will be illuminated before you. Your task is to match the duration of the noise with the duration of the light by turning off the light. The light is shut off by pressing the key. Do not try to estimate by counting or by attending to heartbeats, breathing, etc.

For the alternate task the words "white noise" and "red light" were interchanged as appropriate. In this and the subsequent experiments, 10 observers made two matches for each standard stimulus. Order of standards (light or sound) were counterbalanced, and the stimulus values were irregular and different for each observer.

The results of matching the duration of the light to that of the noise, and vice versa, are shown in Fig. 8. As usual, the observers tended to constrict or shorten the range of their adjustments on whichever variable was under their control, thereby producing a regression effect.

In addition there is a systematic tendency for the data in Fig. 8 to fall below the 45^0 diagonal which would correspond to equal physical durations. On the average, it required less duration of the noise to match the apparent duration of the light. The magnitude of this effect can be measured by averaging the appropriate columns in Table 1. Thus the standard stimulus (light or sound) was, on the average, 1.55 sec. When the duration of the light was adjusted by the observer, its average duration was 1.45 sec. When the duration of the noise was adjusted, its average duration was only 1.13 sec. Other

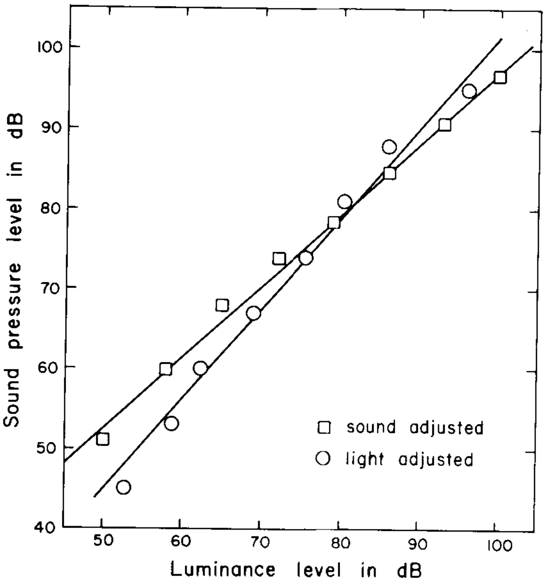

Fig. 7. Cross-modality matches between loudness and brightness. Each point is the decibel average of two adjustments by each of 10 observers.

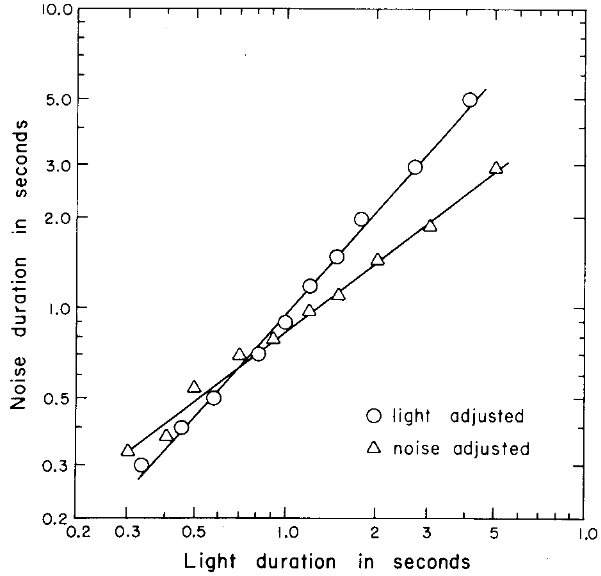

Fig. 8. Matching functions for durations defined by a light and by a noise. Each point is the geometric mean of two adjustments by each of 10 observers.

experimenters have also found that a given duration of sound seems longer than the same duration of light (Behar & Bevan, 1961; Goldfarb & Goldstone, 1964).

For durations marked by light or by sound, counterbalanced number-matching experiments were carried out with 10 observers. In the two magnitude-estimation experiments, no modulus was designated. The first stimulus was a different duration for each observer, and he was told to assign it a number appropriate to its apparent duration. He then assigned numbers proportional to the apparent durations of the other stimuli presented in irregular order. This procedure of using no designated standard modulus has the advantage that it avoids an unnecessary constraint on the behavior of the observer. Furthermore it helps to avoid a piling up of judgments on certain preferred round numbers.

Table 1. Geometric means (in seconds) for the matching of durations marked by light and by sound. The duration of one stimulus was adjusted to match the duration of a given stimulus duration in the other modality. The standard deviations are in logarithmic units. Each entry is based on two matches by each of 10 observers.

	Duration of Dependent Variable			
Criterion	Light adjusted		Noise adjusted	
duration	Geom. mean	SD(log)	Geom. mean	SD(log)
0.3	0.34	0.14	0.33	0.11
0.4	0.45	0.12	0.38	0.13
0.5	0.57	0.09	0.55	0.10
0.7	0.82	0.09	0.70	0.18
0.9	1.00	0.07	0.80	0.18
1.2	1.21	0.11	1.01	0.13
1.5	1.48	0.07	1.14	0.10
2.0	1.82	0.09	1.46	0.08
3.0	2.69	0.08	1.91	0.09
5.0	4.09	0.09	3.00	0.09
Avg. 1.55	1.447	0.95	1.128	1.01

In the two magnitude-production experiments it was thought advisable to employ a standard duration (3 sec.) and to call it 10. The reason for imposing this constraint hinged on the finite capacity of the apparatus to produce time intervals, especially very short time intervals. Even though standards should generally be avoided, the use of a specified standard modulus in matching experiments sometimes helps to keep the observer's adjustments on scale.

Numbers between 1 and 50 and spaced roughly logarithmically were read to the observer in irregular order, and he pressed the appropriate key to produce what he judged to be a proportional duration.

The results of all four experiments are shown in Fig. 9. Whether the interval is marked by light or by sound, the regression effect is clearly present.

In their free matching of numbers to duration, it is evident from Fig. 9 that the observers tended to use smaller numbers to designate the intervals marked by light than by sound. This difference accords with the evidence in Fig. 8 that a given duration of sound seems longer than the same duration of light.

The slopes (exponents) of the power functions in Fig. 9 are 0.87 and 1.20 for noise, and 0.93 and 1.16 for light. The geometric average of all four slopes is 1.03, which suggests that apparent duration is almost a linear but slightly accelerating function of physical duration.

Geometric means and standard deviations are recorded in Table 2. The *relative* variability proved notably constant throughout each experiment, i.e., the variability expressed in log units tends to be constant. The variability of the magnitude estimations is artificially high because the tabled values include the component of variability attributable to the fact that each

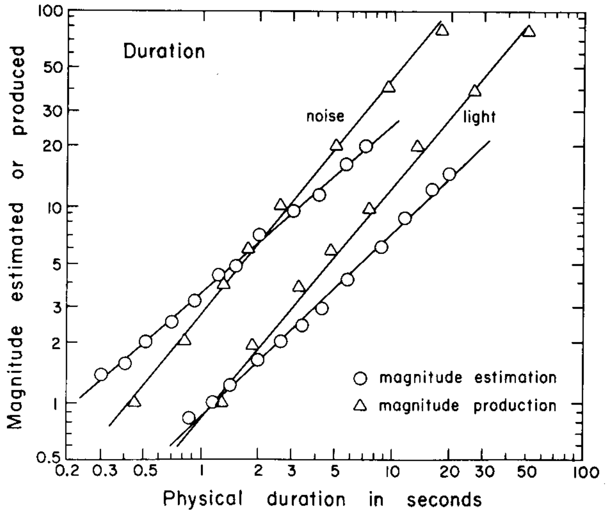

Fig. 9. Matching functions between number and duration. Each point is based on two matching judgments by each of 10 observers. The functions for the intervals marked by a light have been shifted to the right by a factor of 3.

Table 2. Geometric means (in seconds) and standard deviations (in logarithmic units) of the results of magnitude estimation and magnitude productions of duration. Each entry is based on two judgments by each of 10 observers.

	Magnitude Estimation			
Criterion	White noise		Red light	
duration	Geom. mean	SD(log)	Geom. mean	SD(log)
0.3	1.37	0.35	0.86	0.47
0.4	1.58	0.39	1.04	0.47
0.5	2.03	0.39	1.28	0.44
0.7	2.56	0.38	1.68	0.48
0.9	3.28	0·40	2.05	0.43
1.2	4.38	0.34	2.55	0.44
1.5	4.87	0.37	3.09	0.48
2.0	7.08	0.34	4.47	0.42
3.0	9.52	0.36	6.22	0.44
4.0	11.5	0.36	9.05	0.40
5.5	16.4	0.45	12.3	0·42
7.0	20.8	0.29	15.3	0.37

	Magnitude Production			
Criterion	White noise		Red light	
number	Geom. mean	SD(log)	Geom. mean	SD(log)
1.0	0.43	0.17	0.42	0.20
2.0	0.81	0.12	0.68	0.22
4.0	1.29	0.13	1.12	0.15
6.0	1.73	0.11	1.64	0.15
10.0	2.49	0.10	2.67	0.17
20.0	4.92	0.10	4.69	0.17
40.0	9.18	0.18	9.34	0.17
80.0	17.1	0.18	17.6	0.18

observer chose his own modulus. This large component of variability may be removed by correcting each observer's judgments by the factor that makes his mean agree with the grand mean of the group (Lane, Catania, & Stevens, 1961). The procedure is, of course, best carried out in logarithms. The grand mean of all the number estimations by all observers is computed and also the mean for each observer. Then the difference between a given observer's mean and the grand mean is added to each of that observer's log scores. This operation leaves unaltered the slope (exponent) of the magnitude function for each observer, but it minimizes the sum of the squared deviations of his estimates around the group regression line. Since the foregoing procedure has already been illustrated in other contexts (for example, Stevens & Guirao, 1964), the so-called intercept variability (due to the observer's choice of modulus) was not partialed out of the values in Table 2.

DISCUSSION

The variety of examples in Figs. 1-9 suggests that the regression effect may afflict psychophysical judgments of all sorts. The question about its universality may be put concretely. If observers tried to match two tones of identical frequency, say, would the loudness balances show a regression effect? The answer is not known, but the regression effect would presumably be there, even though in any given experiment the difference between the regression lines might be obscured by other factors. As a general principle, two regression lines are produced whenever the correlation between two variables is less than perfect.

What causes the regression effect? Many factors may contribute to the regression tendency, including, ultimately, the irreducible noise that goes with any attempt to measure any empirical value. The regression factors, like errors in general, do not always lend themselves to ready identification and easy extermination. One lesson is plain, however. As we undertake to penetrate the fog of systematic and random errors in order to achieve a closer approximation to the unadulterated value of an exponent, or any other quantity, a single experiment will not suffice. Nor is it likely that statistical analyses can substitute for experimental manipulations, for the assumptions basic to most statistical models are met in the laboratory only under special and unusual circumstances.

Observation of the regression effect in a variety of cross-modality matching judgments has revealed a few of the more obvious factors that may contribute to it.

Incommensurate ranges. If one variable can be varied over a wide range, but the other variable is artificially restricted to a short range, a large regression effect will most likely ensue. Therefore, in setting the standard or criterion values that are to be matched by the observer's manipulation of another variable, the experimenter will want to avoid setting a range of standards that the variable continuum cannot be made to match. The problem arises whenever the range of one variable is more limited by apparatus, or some other factor, than the range of the other variable. In magnitude estimation, for example, the range of numbers that can be used is virtually unlimited, but in magnitude production the range of the variable continuum is often severely limited. In choosing the criterion numbers that are to be matched by magnitude production, the experimenter will try to keep the range of numbers commensurate with the range available on the variable stimulus.

A series of preliminary experiments can help in such decisions. In principle, the problem of adjusting the ranges may be solved by the often neglected but powerful technique of experimental iteration. The ranges could be varied in successive experiments until a minimum regression angle was achieved. Usually, however, reasonably satisfactory ranges will become evident after a few trials.

Observer differences. Observers differ in their approach to a matching task. Some restrict their responses to a cautiously narrow range, others spread their responses more widely. The overly conservative tendency may sometimes produce a remarkable regression angle. If the task is number matching, for example, magnitude estimation gives a function that is much flatter than that of the median observer, but magnitude production gives a function that is correspondingly much steeper.

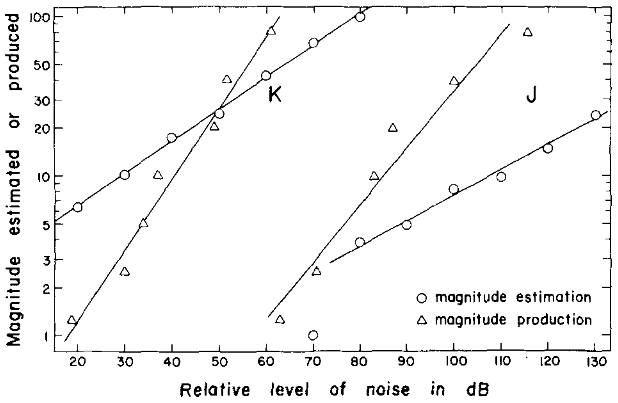

Fig. 10. Matching functions between number and loudness for two individual observers. Each point is the geometric mean of two estimations or productions. The stimulus was a band of noise 500-5000 Hz.

A search through some past experiments for examples of observers whose results gave exceptionally large regression angles produced the two samples shown in Fig. 10. They are from a study by Stevens and Guirao (1962). The stimulus was a band of noise 500-5000 Hz. Written instructions were given for magnitude estimation as follows:

> I am going to present a series of noises. Your task is to tell how loud the noises sound by assigning numbers to them. Call the first noise any number that seems to you appropriate. Your task is to assign numbers proportional to your subjective impression. For example, if a noise sounds 3 times as loud, assign a number 3 times as large as the first. If it sounds 1/5 as loud, assign a number 1/5 as large, and so forth.

Although no standard modulus was designated for magnitude estimation, for the experiment on magnitude production it was thought desirable to assign a modulus in order to keep the adjustments of all the observers within the range of the available stimuli. Consequently, the instructions were as follows:

> I am going to present a noise whose loudness will be called 10. Then I will present a series of numbers, one at a time. Your task is to adjust the noise until its loudness seems proportionate to the numbers I give you, remembering that the first stimulus was called 10. For example, if I say 30 you should adjust the noise so that it seems 3 times as loud, and so forth. When adjusting the noise, it is helpful to bracket the desired value by approaching it from above and below.

Several interesting features are evident in Fig. 10. Note that observer K chose to use a larger modulus than J, whose choice of modulus was such that the faintest stimulus was called one, a value quite far from the line described by J's other estimates. This and other "round number" tendencies sometimes distort a function, but they can be partially averaged out if each observer chooses a different modulus. Thus K's estimate for the faintest noise is consistent with his other estimates.

Since a modulus was designated for the magnitude productions, the sound pressure levels that the observer produced centered about a common value. That, in fact, was the purpose of the assigned modulus.

In their magnitude productions, the two observers J and K produced fainter noises, on the average, than were used in the experiment on magnitude estimation. Perhaps they found the louder noises disagreeable, a factor that could have contributed to their large regression effects.

Observers J and K, it should be noted, had participated in other similar experiments, always, of course, without knowledge of the outcome in terms of regression angles. It is not clear that experience as an observer has much if any effect on regression.

Although each point in Fig. 10 represents the geometric mean of only two matching judgments, other studies suggest that, if more matching judgments had been made, the results would have been little altered. Observers settle quickly into a pattern or groove and they tend thereafter to give self-consistent judgments, unless, of course, with long repeated testing they become bored and lose interest. In a given experiment, the increment of useful new information decreases quite rapidly after the first series of matching judgments. Brief experiments yield relatively more information than protracted ones.

The regression effect is not greatly dependent on the observer's knowing the range of stimuli with which he will be presented. In the matching experiments reviewed above, the observers were given no advance information concerning the stimulus range. Yet the regression effect was typically evident in the first few judgments that the observers made. As a matter of fact, the regression effect was seen in the group results of an experiment in which each observer made only one matching judgment (Stevens & Poulton, 1956). The elimination of the context provided by the other stimuli did not seem to abolish the regression effect.

One of the more obvious indications of Fig. 10 is that neither magnitude estimation nor magnitude production gives a good measure of the loudness exponents for the two observers. Thus K's function has an exponent of

0.4 by magnitude estimation and 0.9 by magnitude production. The geometric mean of those two values is 0.6, which happens to be the exponent of the loudness function recommended for engineering purposes by the International Standards Organization.

But the value 0.6 may or may not prove to be the best measure of K's exponent. The problem of determining an unbiased psychophysical matching function for a single observer cannot be solved by number matching alone. We should also want to see what happens under other cross-modality matching procedures. If we are seriously interested in the power function for a particular person, we will want to know the results of a balanced array of cross-modality matching tasks. Since some continua may not lend themselves to manipulation by the observer, the battery of cross-modality matching functions cannot always be balanced in terms of the regression effect.

Averaging. Can the regression effect be averaged out? The answer to that question is probably yes, although major uncertainties remain open. As a procedure for averaging the two regression lines, an argument in favor of the geometric mean of the two slopes is given by Indow and Stevens (1966). The rationale is that a representative line synthesizing the two regression lines ought to remain the same when the ordinate and abscissa are interchanged. Only the geometric mean provides the desired invariance.

Not all questions concerning averaging can be disposed of so easily, however, for there remains a question whether the factors that produce the regression effect in a given experiment have been affected symmetrically by interchanging the variable stimuli. It seems highly plausible that the most representative function lies somewhere between the two regression lines, but does it coincide with the geometric mean, or does it lie closer to one or the other of the two lines? Otherwise said, the distorting constraints that affect one variable, say, magnitude estimation, may be different from the constraints that affect the other variable, which may involve, for example, the turning of a knob that has peculiar characteristics of its own. It seems unlikely that the errors in two such different tasks as magnitude estimation and magnitude production would inevitably cancel exactly and completely. That problem, like many other issues concerning the nature of systematic errors and anomalies, remains a continuing challenge.

It is nevertheless apparent that the use of magnitude estimation alone, without a counterbalancing experiment involving magnitude production, tends in general to underestimate the exponent of the power function. In the typical experiment involving a group of observers, the method of magnitude estimation may provide an adequate lower bound on the value of the exponent, but an estimate of the upper bound calls for magnitude production, or some analogous procedure. As mentioned above, a direct determination of the probable upper bound may not prove readily possible with some kinds of stimuli.

References

Aiba, T. S., & Stevens, S. S. Relation of brightness to duration and luminance under light- and dark-adaptation. *Vision Res.*, 1964, 4, 391-401.

Behar, I., & Bevan, W. The perceived duration of auditory and visual intervals: cross-modality comparisons and interactions. *Amer. J. Psychol.*, 1961, 74, 17-26.

Goldfarb, J. L., & Goldstone, S. Properties of sound and the audio-visual differences in time judgment. *Percept. mot. Skills*, 1964, 19, 606.

Hollingworth, H. L. The central tendency of judgment. *J. Phil. Psychol.*, 1910, 7, 461-469.

Indow, T., & Stevens, S. S. Scaling of saturation and hue. *Percept. & Psychophys.*, 1966, 1, 253-271.

Lane, H. L., Catania, A. C., & Stevens, S. S. Voice level: Autophonic scale, perceived loudness, and effects of sidetone. *J. Acoust. Soc. Amer.*, 1961, 33, 160-167.

Reynolds, G. S., & Stevens, S. S. Binaural summation of loudness. *J. Acoust. Soc. Amer.*, 1960, 32, 1337-1344.

Stevens, J. C., & Marks, L. E. Cross-modality matching of brightness and loudness. *Proc. Nat. Acad. Sci.*, 1965, 54, 407-411.

Stevens, S. S. The measurement of loudness. *J. Acoust. Soc. Amer.*, 1955, 27, 815-829.

Stevens, S. S. Calculation of the loudness of complex noise. *J. Acoust. Soc. Amer.*, 1956, 28, 807-832.

Stevens, S. S. On the operation known as judgment. *Amer. Sci.*, (in press).

Stevens, S. S., & Guirao, Miguelina. Loudness, reciprocality, and partition scales. *J. Acoust. Soc. Amer.*, 1962, 34, 1466-1471.

Stevens, S. S., & Guirao, Miguelina. Scaling of apparent viscosity. *Science*, 1964, 144, 1157-1158.

Stevens, S. S., & Poulton, E. C. The estimation of loudness by unpracticed observers. *J. exp. Psychol.*, 1956, 51, 71-78.

Zwicker, E. Über psychologische und methodische Grundlagen der Lautheit. *Acustica* 8, *Akus. Beih.*, 1958, 1, 237-258.

Notes

1. Research supported in part by the National Science Foundation Grant GB-3211, and in part by the National Institutes of Health Grant NB-02974 and Public Health Postdoctoral Fellowship No. 7-F2-MH-19, 534-02 (Laboratory of Psychophysics Report PPR-330).

2. Present address: Psychology Services, Veterans Administration Hospital, Northampton, Massachusetts.

(Accepted for publication October 14, 1966.)