Malleability of "ratio" judgments of occupational prestige

CURTIS HARDIN and MICHAEL H. BIRNBAUM
California State University, Fullerton

Subjects judged "ratios" and "differences" of the prestige of occupations. "Ratio" and "difference" judgments were nearly monotonically related, consistent with the hypothesis that subjects used the same operation for both comparisons. To test the invariance of "ratio" judgments, different groups were exposed to different numerical examples of "ratios." When the largest "ratio" mentioned in the instructions was "4," the median and modal "ratio" of the prestige of *physician* relative to *trash collector* was 4. When the largest "ratio" mentioned was "64," the median and modal "ratio" judgment for *physician* relative to *trash collector* was 64. This malleability of "ratio" judgments compromises the ratio model of stimulus comparison, but is consistent with Birnbaum's theory that "ratio" judgments are governed by subtraction on the same interval scale that underlies "difference" judgments.

A persistent and troubling problem for the foundation of a quantitative science of psychology has been the finding that the results of different methods for the measurement of subjective value do not agree. In a recent volume of contributions on social attitudes and psychophysical measurement (Wegener, 1982b), a number of authors discussed the scaling of the prestige of occupations. These authors agreed that the scaling of prestige behaves like the scaling of other psychological and psychophysical continua: Different methods of scaling lead to different scales of prestige (Cross, 1982; Dawson, 1982; Schneider, 1982; Wegener, 1982a).

For example, Wegener (1982a) reported a study of 1,796 subjects who evaluated 16 occupations by category ratings and "magnitude" methods of scaling. Consistent with the typical finding, Wegener found that the two scales were not linearly related. Scales based on magnitude estimations or "ratio" judgments and ratio models tend to be a nonlinear, positively accelerated function of scales based on category ratings, "difference" judgments, or scales based upon subtractive models applied to paired comparisons (Dawson & Brinker, 1971; Kuennapas & Wikstroem, 1963: Stevens, 1966a; Treiman, 1977).[1]

Birnbaum and Veit (1974) noted that many of the failures to find agreements of scales in the literature can be attributed to two factors:

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 1990, Vol. 103, No. 1, pp. 1–20

(a) experimenters (erroneously) assumed that subjects performed whatever operations they were instructed to perform, and (b) experimenters did not utilize experimental designs and analytic methods that allowed for the separation of response "biases" and contextual effects from comparison processes. Birnbaum and Veit found that "ratios" and "differences" were monotonically related, consistent with Torgerson's (1961) suggestion that perhaps it is the theoreticians and experimenters, rather than the subjects, who are responsible for the contradictions in the scaling literature. Perhaps subjects do the same thing, whether instructed to judge "ratios" or "differences," but experimenters have interpreted the data differently.

Birnbaum (1978, 1980, 1982) noted that judgments of psychological "ratios" and "differences" are monotonically related for a number of continua including heaviness, loudness, pitch, darkness, and likableness of adjectives. If subjects in these studies were actually using both ratio and difference operations, judgments of "ratios" and "differences" would not be monotonically related, but would instead show a distinct pattern of interrelationships (Birnbaum, 1980; Krantz, Luce, Suppes, & Tversky, 1971; Miyamoto, 1983). However, results from many studies are compatible with the hypothesis that subjects use the same operation for both tasks, despite the different instructions (Birnbaum, 1980, 1982, in press). For example, Schneider (1982) found that judgments of "ratios" and "differences" of occupational prestige lead to nearly identical scaling solutions when fit to the Euclidean model, because subjects seem to produce the same rank order whether instructed to compare prestige by a "ratio" or a "difference" operation.

If it is concluded that subjects are, in fact, using the same operation for "ratios" and "differences," the problem of deciding whether that single operation is best represented as a ratio, a difference, or something else still remains. To decide this issue, it is necessary to provide additional theoretical and experimental constraints, such as the four-stimulus comparison experiments involving "ratios of differences" and "differences of differences" as well as scale convergence across a wider realm of results. Results from experiments based on these constraints have supported subtractive theory over ratio theory (Birnbaum, 1978, 1979, 1980, 1982, in press).

In the case of occupational prestige, however, arguments have been advanced to prefer ratio theory over subtractive theory. Kuennapas and Wikstroem (1963) noted that magnitude estimation, "ratio" estimations, and paired comparisons of occupational prestige would all be mutually consistent if Case V of Thurstone's law, a subtractive model in which each stimulus has an equal dispersion (Thurstone, 1927; Torgerson, 1958), were replaced by the Case VI assumption that stimulus dispersions vary in proportion to stimulus value. Additionally, Stevens (1966a) noted that the "ratio" of the most extreme comparison was roughly constant (about 30:1) in the studies he reviewed. Recently, Cross (1982) further suggested that magnitude and cross-modality matching methods are consistent with the assumption that ratios of magnitude are independent of the method of measurement. For example, in the studies cited by Cross (1982), the prestige of a *physician* is roughly 1.7 times greater than that of an *engineer*, whether scaled by loudness, brightness, or magnitude estimation. If empirically supported, "ratio" invariance would strengthen the viability of ratio theory; indeed, subtractive theory would not, in general, predict such consistency.

In subtractive theory, the comparison of two stimuli on an interval scale is a directed distance that has an arbitrary unit. According to Birnbaum's (1978, 1980, 1982) one-operation theory, "ratio" judgments should be malleable, because they are merely a judgment transformation of arbitrarily scaled differences. This theory can be represented as follows:

$$R_{ij} = J_R(s_j - s_i) \tag{1}$$

$$D_{ij} = J_D(s_j - s_i) \tag{2}$$

where R_{ij} and D_{ij} are the judged "ratio" and "difference" between stimuli j and i, respectively; J_R and J_D are the strictly monotonic judgment functions; and the scale values of the stimuli, s_j and s_i, are compared by subtraction in both cases. It is assumed that the scale values are the same in both equations, though the J functions are not necessarily equivalent.

In two-operation theory, the subtractive operation in Equation 1 is replaced by a ratio as follows:

$$R_{ij} = J_R(s_j/s_i) \tag{3}$$

where the terms are defined as above.

One-operation theory (Equations 1 and 2) implies that "ratios" and "differences" will be monotonically related. However, two-operation theory (Equations 2 and 3) implies that there is no function that will assign "ratios" to "differences" if both s_i and s_j are independently manipulated (see Birnbaum, 1980). For example, $2 - 1 = 3 - 2$, but $2/1 > 3/2$; $2/1 = 4/2$, but $2 - 1 < 4 - 2$; and $2/1 > 5/3$, but $2 - 1 < 5 - 3$.

In the present experiments, subjects were asked to judge both "ratios" and "differences" of prestige to test the one-operation versus two-operation theories of stimulus comparison. The major purpose was to test the "ratio" invariance prediction of ratio theory against the implications of the judgment function of subtractive theory. We used the method of Mellers, Davis, and Birnbaum (1984), and mentioned different example "ratios" incidentally in the instructions. According to ratio theory, if "ratio" judgments are to be taken at face value, such a manipulation should have no effect, or at worst, the manipulation should have the effect of a linear transformation of "ratio" judgments (Gulliksen, 1959).

However, according to Birnbaum's (1978) subtractive theory, "ratios" are arbitrary numbers, and numerical "ratios" cannot be safely exported from one situation to another without a theory of the judgment function. When the examples form a geometric series, "ratio" judgments are predicted to be an exponential function of subjective differences. If subjects use "ratios" of 1/4, 1/2, 1, 2, and 4 for a set of equally spaced categories, these "ratios" are exponentially related to the categories. Therefore, subtractive theory predicts that changing the examples will change the J_R function (Birnbaum, 1980, 1982), so "ratio" judgments need not be linearly related to each other.

EXPERIMENT

METHOD

Subjects were presented with pairs of occupations, and for each pair, they estimated the "ratio" of the prestige of the first occupation to the second occupation. In a separate series of trials, they also judged the "difference" in prestige for each pair. Half of the subjects made "ratio" judgments first, and half judged "differences" first. In addition, two groups of subjects received different example responses for the "ratio" task.

Instructions

The "difference" instructions requested subjects to estimate the "difference in prestige between the first and the second occupations." Occupational prestige "differences" were rated on a 13-point scale with category labels varying from "6 = *The first occupation is extremely more prestigious than the second occupation*" to "−6 = *The first occupation is extremely less prestigious than the second occupation*"; "0" was "*equal*."

The "ratio" instructions requested subjects to estimate the "ratio of the prestige of the first occupation to the second occupation." For both "ratio" conditions, the modulus was "100 = *The first occupation is just as prestigious as the second occupation*."

Half of the subjects ("ratio-4" condition) received a page that illustrated "ratio" judgments with examples as follows: "400 = *The first occupation is four times as prestigious as the second occupation*"; "200 = *The first occupation is* *two times as prestigious as the second occupation*"; "100 = *The first occupation is just as prestigious as the second occupation*"; "50 = *The first occupation is one half as prestigious as the second occupation*"; and "25 = *The first occupation is one fourth as prestigious as the second occupation.*"

The other half of the subjects ("ratio-64" condition) received examples as follows: "6400 = *The first occupation is sixty-four times as prestigious as the second occupation*"; "800 = *The first occupation is eight times as prestigious as the second occupation*"; "100 = *The first occupation is just as prestigious as the second occupation*"; "12.5 = *The first occupation is one eighth as prestigious as the second occupation*"; and "1.56 = *The first occupation is one sixty-fourth as prestigious as the second occupation.*" For this group, numerical examples of "3200," "1600," "400," "200," "50," "25," "6.25," and "3.12" were also printed on the page between the written examples without written interpretations. Note that this manipulation should be purely incidental if numerical "ratio" judgments are to be taken at face value, because in both cases the same subjective ratio corresponds to the same numerical response.

For all tasks, subjects were encouraged to use whatever responses they thought appropriate, including values not mentioned.

Stimuli and design

Twelve occupations were used in a 7 × 7, First Occupation × Second Occupation factorial design. First occupations were: *trash collector*, *car mechanic*, *plumber*, *high school teacher*, *nurse*, *college professor*, and *physician* (Treiman's [1977, Appendix A] values for these occupations are 13, 43, 34, 64, 54, 78, and 78, respectively). Second occupations were: *trash collector*, *factory worker*, *carpenter*, *secretary*, *police officer*, *architect*, and *physician* (Treiman's values are 13, 29, 37, 53, 40, 72, and 78, respectively). *Trash collector* and *physician* were used for both factors to check the subject's understanding of the instructions and to allow both sets of occupations to be calibrated on the same scale.

Procedure

Subjects received booklets containing computer-generated, randomly ordered, stimulus pairs with both "ratio" and "difference" instructions. Half of the booklets asked for "difference" judgments first, and half asked for "ratio" judgments first. Also, half of the subjects were randomly assigned to one of the two "ratio" conditions. Packets contained detailed instructions as well as practice trials for each of the tasks. The experimenter checked the practice trials to ensure that subjects understood the instructions. All subjects showed at least superficial understanding of the tasks by judging the "differences" of *physician* to *physician* or *trash collector* to *trash collector* as 0, and the "ratios" as 100 (1 × 100). Subjects then completed the 49 experimental trials at their own pace before proceeding to the next task. All subjects completed the experiment within the 50-min session.

Subjects

Participants were 50 undergraduates at California State University, Fullerton, who received extra credit in an introductory psychology course.

RESULTS

Figures 1 and 2 plot geometric mean "ratio" judgments against mean "difference" judgments, using a separate figure for each of the "ratio" conditions. Within each figure, a separate type of symbol is used for each level of the second occupation. One-operation theory (Equations 1 and 2) implies that "ratios" and "differences" should be monotonically related. On the other hand, two-operation theory (Equations 2 and 3) implies that "ratios" and "differences" should not be related by a function of a single variable but should form a fan of intersecting lines, one for each level of the second occupation, in which "ratios" corresponding to equal "differences" should be unequal (Birnbaum, 1980, Fig. 3; for example, $3/1 > 4/2$ although $3 - 1 = 4 - 2$). The data in Figures 1 and 2 indicate that "ratio" judgments are very nearly exponential functions of "difference" judgments. Therefore, these data appear consistent with the hypothesis that subjects used the same operation to compare stimuli in both tasks, contrary to the instructions.

The manipulation of the example "ratios" had a huge effect. When the largest example "ratio" was "4," the occupation of *physician* was

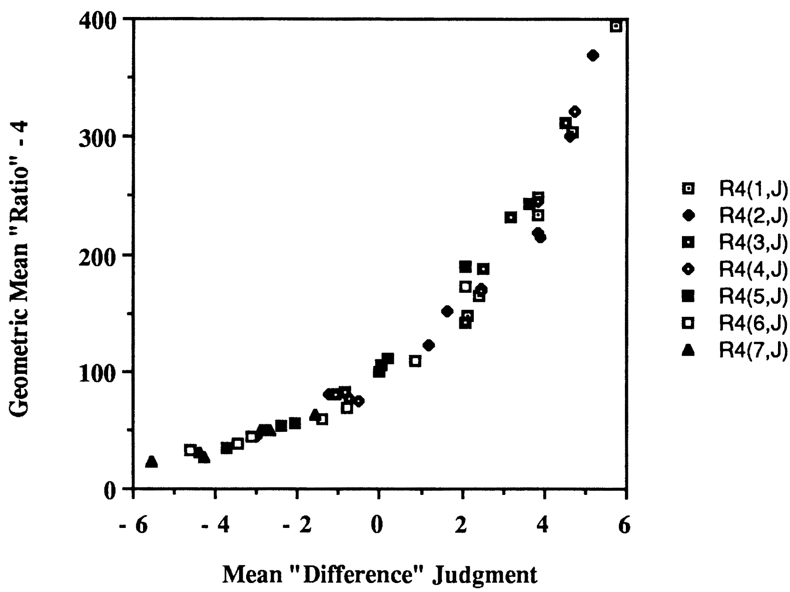

Figure 1. Geometric mean "ratio" judgments plotted against mean "difference" judgments, with a separate point for each pair of occupations (in this condition, the largest example of a "ratio" was "four times"; different symbols are used for different levels of the second occupation)

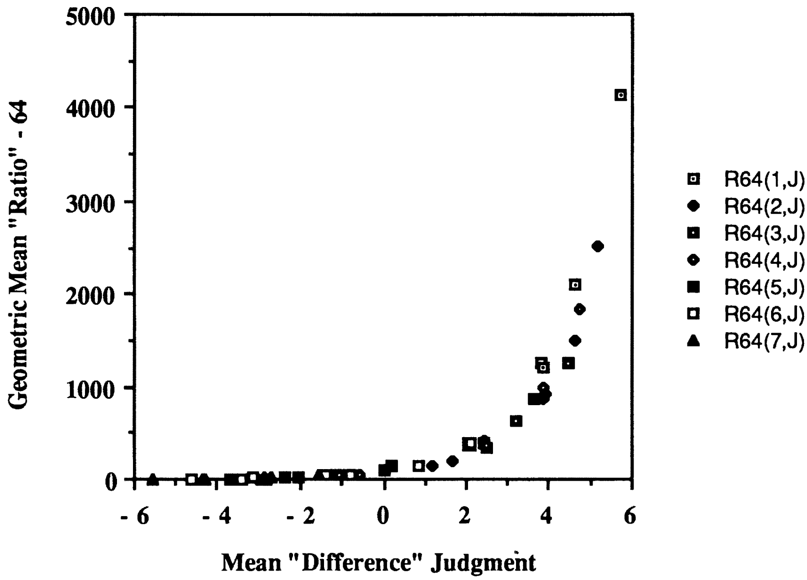

Figure 2. Geometric mean "ratio" judgments plotted against mean "difference" judgments as in Figure 1, except that the largest example of a "ratio" was "64 times"

judged to be "4 times" more prestigious than the occupation of *trash collector* (mode = median = 4; mean = 4.05; geometric mean = 3.95). Only 3 subjects (of 26) gave "ratios" greater than "4," and the largest was "8." However, when the largest example was "64," *physician* was judged "64 times" more prestigious than *trash collector* (mode = median = 64; mean = 49.17; geometric mean = 41.40). In the "ratio-64" condition, only 2 subjects (of 24) gave a "ratio" less than "16," 13 subjects used "64," and one said "65."[2]

Similar results were obtained for "ratios" less than "1." When the example responses ranged from 1.56 ("1/64") to 6400 ("64"), 13 of the 24 subjects used this exact response range, and only one subject used responses more extreme (from 1.45 to 6500). When the examples ranged from 25 ("1/4") to 400 ("4"), 19 of 26 subjects used this exact range.

Figure 3 plots judged "ratios" in the "64" condition versus "ratios" obtained in the "4" condition ("64" indicates that the incidental examples ranged from "1/64th" to "64 times"; and "4" indicates that the incidental examples ranged from "1/4th" to "4 times"). Figure 3 shows that "ratio" judgments are not linearly related to one another. Instead, both sets of "ratio" judgments are roughly expo-

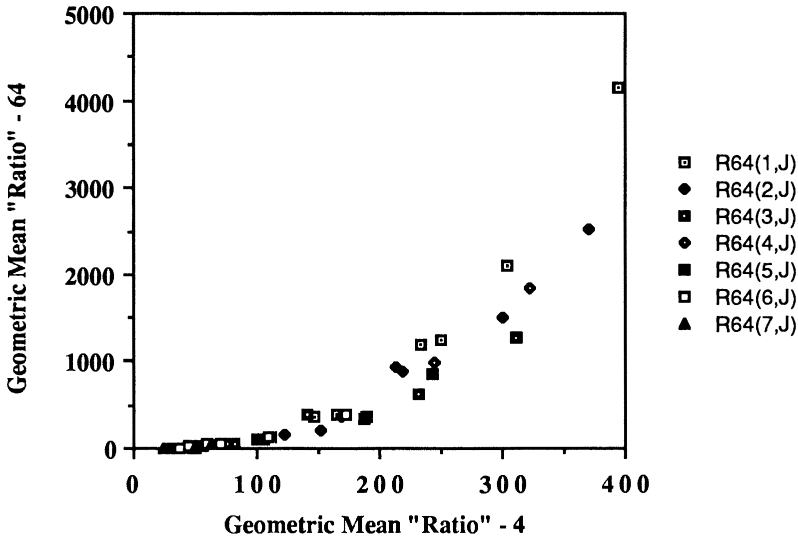

Figure 3. Geometric mean "ratio" judgments from Group "64," plotted against geometric mean "ratio" judgments from Group "4" (different symbols are used for different levels of the second occupation, as in Figure 1)

nentially related to "difference" judgments, and this change in examples appears to act as a multiplier, prior to the exponential transformation; therefore, median "ratio" responses in the "64" condition are roughly the cube of the corresponding values in the "4" condition. Put more simply, subjects seem to use so-called "ratios" as they use category judgments, contrary to the notion that the numerical responses can be interpreted as true ratios.

Figure 4 plots the mean "difference" judgments, averaged over both groups of subjects, plotted as a function of the column marginal means for the first stimulus. If the J_D function were linear, Equation 2 would imply that the curves in Figure 4 should be linear and parallel.[3] The data are almost parallel, but show a small systematic departure consistent with the hypothesis that J_D is steeper in the region of "no difference." The interaction contains 2% of the systematic variance, but is significant, $F(36, 1764) = 12.25$.

Figures 5 and 6 plot geometric mean "ratio" judgments, in the same format as Figure 4, for the two "ratio" conditions. According to one-operation theory (Equations 1 and 2), if J_R is an exponential function, the data should form a bilinear, divergent fan of straight lines that intersect at a common point.[4] Again, the data in each case approximate these predictions, though again there is a slight depar-

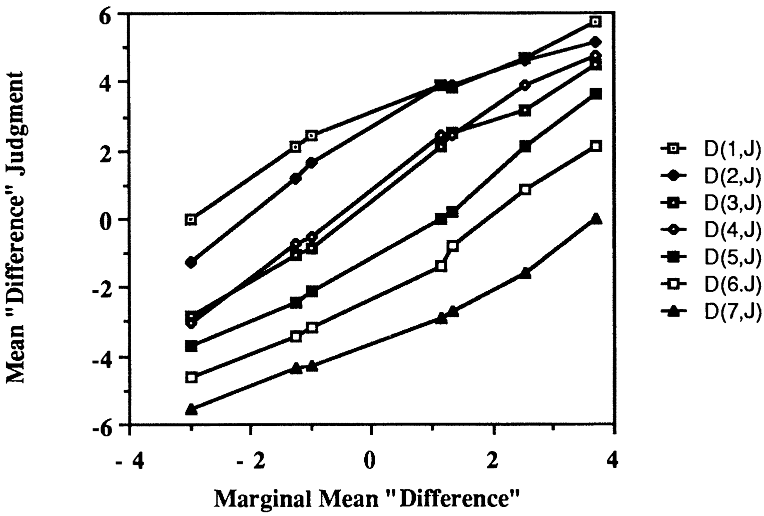

Figure 4. Mean "difference" judgments plotted against marginal mean "difference" judgments for the first occupation (averaged over the second occupation), with a separate curve for each level of the second occupation

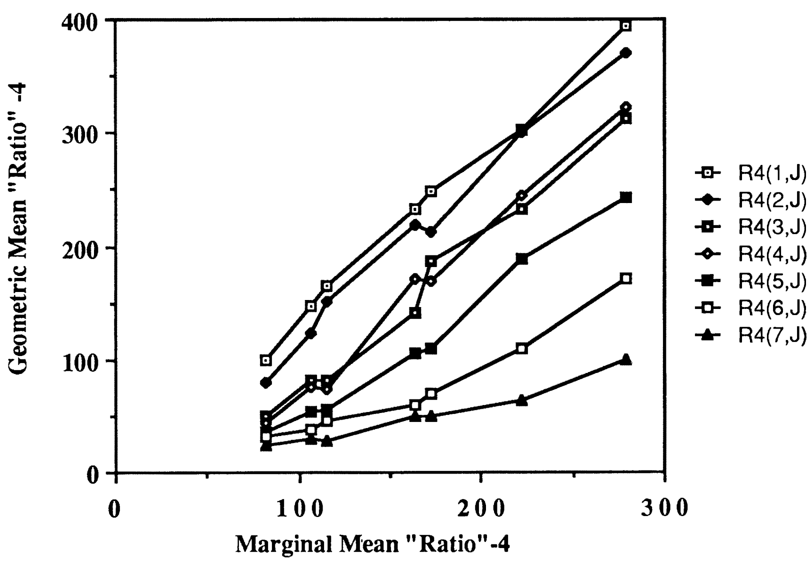

Figure 5. Geometric mean "ratio" judgments from Group "4," plotted as in Figure 4

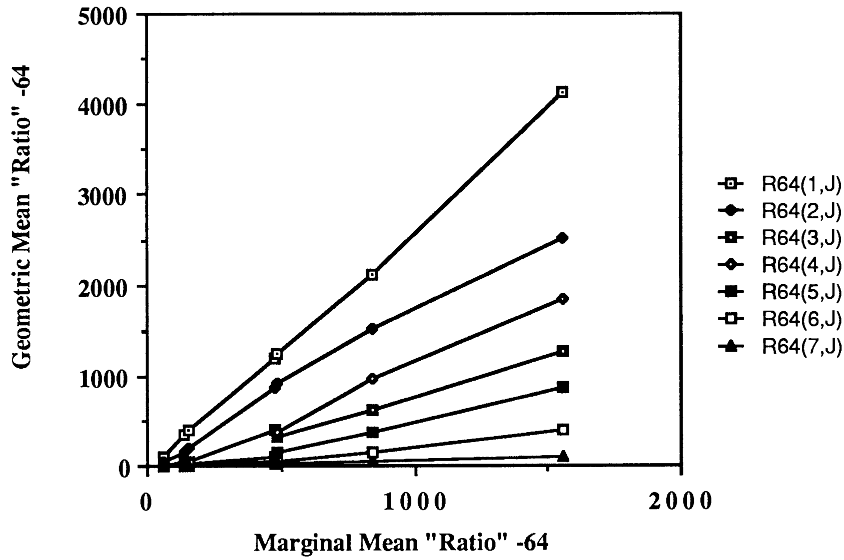

Figure 6. Geometric mean "ratio" judgments from Group "64," plotted as in Figure 4

ture, indicating exaggeration of small differences. The interactions in the logs of the "ratios" for the two conditions are small (1.7% and 2.3% of the systematic variance) but significant, $F(36, 900) = 4.76$, and $F(36, 828) = 3.57$, respectively. This pattern has been found in previous studies (e.g., Elmasian & Birnbaum, 1984), and may be attributed to the judgment functions.

Figure 7 plots the marginal mean logs of "ratio" judgments for the first stimulus as a function of marginal mean "differences." As explained in Notes 3 and 4, these curves should be linear if J_R and J_D in Equations 1 and 2 are exponential and linear, respectively (Birnbaum, 1980). However, if "ratio" judgments are a power function of subjective ratios, according to two-operation theory, then the curves in Figure 7 should be logarithmic, rather than linear. The linearity in Figure 7 indicates that although "ratios" are not linearly related to "differences" (Figures 1 and 2), nor linearly related to each other (Figure 3), there is metric information in the data that can be described by Equations 1 and 2. Linearity of the curves in Figure 7 is compatible with one-operation theory, scale convergence, and the simplifying assumptions that J_D and J_R are linear and exponential, respectively.

Table 1 lists scale values for the 12 occupations, estimated according to the simplified one-operation subtractive theory applied separately to "differences," "ratio-4," and "ratio-64." The simplified theory as-

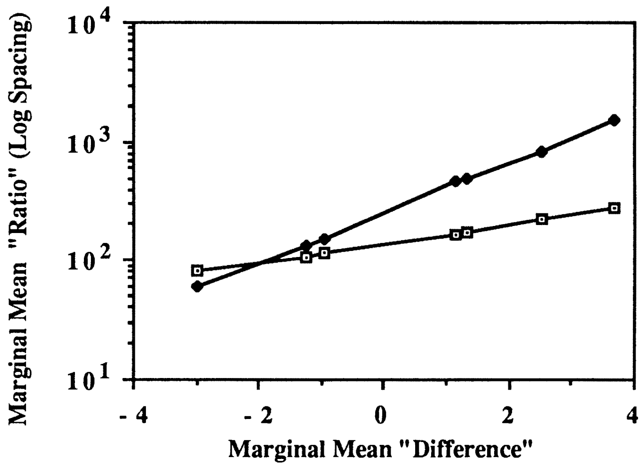

Figure 7. Marginal mean "ratio" judgments (Log spaced) plotted against marginal mean "difference" judgments (open squares depict results for "ratio-4" group; closed symbols show results for "ratio-64" group; linearity of these graphs is consistent with the one-operation theory)

sumes that J_R and J_D in Equations 1 and 2 are exponential and linear, respectively, with arbitrary constants (Birnbaum, 1980). Agreement of the estimated scale values in Table 1 indicates that, despite the nonlinear relationships among the three sets of data, the results can be reconciled by one-operation theory, with a single scale of prestige.

DISCUSSION

The present results appear consistent with the theory that subjects compare prestige of occupations by subtraction, whether instructed to judge "ratios" or "differences." This theory explains why "ratios" and "differences" are monotonically related, as well as why "ratio" judgments are so malleable.

The present results for occupational prestige are entirely compatible with previous studies of "ratio" and "difference" judgments (Birnbaum, 1978, 1980, 1982, in press; DeGraaf & Frijters, 1988; Mellers & Birnbaum, 1982; Schneider, 1982; Veit, 1978). For example, Mellers et al. (1984) found that "ratios" and "differences" of the heaviness of lifted weights are monotonically related; they also found that "ratio" judgments are malleable. By changing the examples, subjects were willing to judge the "ratio" of the same pair of weights to be either "4," "8," or "32," depending upon example "ratios" mentioned in the instructions.

Table 1. Estimated scale values, based on subtractive theory applied separately to "differences" and each "ratio" condition

Occupation	"Differences"	"Ratios" (64)	"Ratios" (4)
Physician	10.0	10.0	10.0
College professor	8.3	8.3	8.1
Architect	7.5	7.7	7.6
Nurse	6.5	6.6	6.1
High school teacher	6.2	6.4	5.7
Police officer	5.7	5.9	5.2
Carpenter	3.4	4.1	3.0
Secretary	3.4	3.8	3.1
Plumber	3.0	3.2	2.6
Car mechanic	2.6	2.6	2.3
Factory worker	0.8	1.4	0.7
Trash collector	0.0	0.0	0.0

Note. Scale values for the two extreme stimuli were arbitrarily scaled to 0 and 10, respectively.

The malleability of "ratios" poses great difficulty for the theory that subjects actually use a ratio operation when judging "ratios" of prestige. Ratio theory would need to develop an explanation of why "ratio" judgments are not linearly related to each other when different examples are used to illustrate the task (Figure 3). If a ratio model is assumed, what is the true ratio of the prestige of *physician* relative to *trash collector*? Subjects are willing to call this "ratio" either "4" or "64," depending upon the examples used. If ratio theory is retained, these results force the theoretician to concede that the ratio model does *not* define ratios of subjective value; the ratio model leads to scales that are only unique to a power transformation.

Why, then, has there been such an apparent consistency of "ratios" in previous studies, as noted by Stevens (1966a) and by Cross (1982)? The answer of subtractive theory is that the judgment functions, J_R and J_D, depend lawfully on the stimulus and response distributions. In other words, the J functions are determined by the context (Mellers & Birnbaum, 1982). When the context is fixed (i.e., the stimulus and response distributions are the same), it is possible to predict from any two stimulus-response functions to the third; and many different theories can then yield successful approximations of the relationship. However, when the stimulus or response distributions, or both, are varied, such properties as "ratio" invariance will be systematically violated (Mellers & Birnbaum, 1982). There has been (apparent) consistency of "ratios" in previous research on occupational prestige only when these distributions were fixed.

Teghtsoonian (1973) noted that across a variety of different continua, subjects tend to use approximately the same response range for magnitude estimation. When the response range is not explicitly manipulated, subjects apparently choose different ranges that average out across subjects to be roughly equal for different dimensions. Within a continuum, manipulation of the stimulus range appears to increase the subject's response range, but not enough to maintain an invariant stimulus-response function. In the present experiment, the stimulus distribution was held constant, but the *response* distribution (the range of examples) was manipulated, with the result that the majority of subjects simply accepted the incidental examples as a category scale. As Birnbaum (1982) noted, in magnitude estimation and cross-modality matching, the stimulus-response function depends upon two factors that the experimenter controls, the stimulus and response ranges.

Subtractive theory gives a coherent account of both the present data and the previous data that seemed so consistent to Stevens (1966a) and Cross (1982). If the subtractive theory is fit to "ratio" judgments, then Case V of Thurstone's law, in which each stimulus has an equal dispersion, can be retained for paired comparisons in Kuennapas and Wikstroem (1963). The other "ratio" results can also be reconciled with subtractive theory by taking a logarithmic transformation of "ratio" judgments obtained when the examples form a geometric series. However, in some studies, other stimulus and response distributions have been used, and in those cases, the J_R function may not be exponential (Mellers & Birnbaum, 1982).

Theories of magnitude estimation and "ratio" judgment

To clarify further how the present approach utilizes the concept of a judgment function (J_R and J_D in Equations 1 and 2), it is useful to contrast the concept of a judgment function with two other simple theories of magnitude estimation (Krueger, 1989):

1. Face-value theory holds that numerical "ratio" judgments can be taken at face value. This theory assumes that the comparison operation is a ratio and that numerical responses reveal those ratios directly without transformation. Stevens (1956, 1966a, 1966b) accepted "ratio" judgments at face value, and he recognized that without such an assumption, magnitude estimations and the ratio model fail to define the exponent in the power law.

2. Psychophysics of numbers theory contends that magnitude estimations must be transformed from the physical scale of numbers to the psychological magnitudes of those numbers (Attneave, 1962; Ekman & Sjoberg, 1965; Krueger, 1989; Rule & Curtis, 1982). Rule

and Curtis have tested and developed this theory extensively and concluded that the psychophysical function for numbers can be approximated as a power function with an exponent of approximately .63. Therefore, according to this theory, a "ratio" judgment of 7 corresponds to a psychological ratio of 3.76 ($3.76 = 7^{.63}$). This theory retains the ratio operation for comparisons, but it allows transformations of the responses in order to explain, for example, why magnitude estimations of "differences" do not fit the subtractive model.

Cross (1982) contrasted the psychophysics of number theory with the face value theory and defended the position that the face value hypothesis should be retained as long as it can be. In his discussion of the psychophysics of numbers theory, he wrote:

> When an experimental finding indicates . . . that the years spent memorizing the multiplication tables and learning how to work with units of measurement have not given us a firm concept of numerical relation—it might be appropriate to question the validity of the assumptions of the experimenter whose criteria for assessing subjective magnitudes lead to such violations of common sense. (p. 87)

Although Cross was arguing against the psychophysics of numbers theory in this passage, the experimental results that cause the "violations of common sense" are only violations to those who assume that subjects are actually judging ratios when so instructed. Such results fall nicely into place when "ratio" judgments are represented as a judgment function of subjective differences.

3. Judgment function theory postulates that the function relating subjective value to numerical responses depends lawfully on the stimulus and response distributions and can be approximated by the principles identified in Parducci's (1982) range-frequency theory, together with some additional assumptions concerning magnitude estimation (Mellers & Birnbaum, 1982). The judgment function describes the way in which subjects assign responses to subjective values: Subjects tend to choose responses such that each response has the same range-frequency value in the response distribution as the stimulus has in the stimulus distribution. A key idea for magnitude estimation is that the judgment function can be exponential if the examples used to illustrate the scale are geometrically spaced (Birnbaum, 1978, 1980). Previous failures to fit ratio models may have occurred because the response examples used were not spaced systematically in those studies.

Some authors (e.g., Marks, 1974; see Krueger, 1989, for a review) have concluded that magnitude estimations are a power function rather than an exponential function of ratings. It is conceivable that there are conditions (choices of examples) for which the power function will provide as good an approximation for this relationship as anything else. However, because this relationship can be manipulated by many factors, any theory that considers it to be invariant faces difficulty (Birnbaum, 1980, 1982; Mellers, 1983a, 1983b; Mellers & Birnbaum, 1982; Mellers et al., 1984). There are conditions under which the power function fails; therefore, it does not seem promising to attempt to force this assumption on the development of future theory.

The theoretical concept of a judgment function offers a better framework for understanding these contextual effects than the theory that magnitude estimations can be corrected by the inverse of the psychophysical function for number. The judgment function is useful for connecting ratings, "ratio" estimations, and other responses within one coherent theory, whereas the number notion needs to explain why a 1–100 rating scale does not involve numbers. The judgment function explains why numerical responses given in rating or magnitude tasks can be so easily manipulated by variations of the stimulus and response distributions, while the psychophysics of number theory does not easily accommodate contextual effects (Birnbaum, 1980; Mellers & Birnbaum, 1982; Mellers et al., 1984).

The sensitivity of the judgment function to changes in the response examples also helps to explain how the judgment function can be exponential for magnitude estimation. When the examples are geometrically spaced (e.g., "1/4," "1/2," "1," "2," and "4"), then the judgment function will be exponential if the subjects treat these responses as equally spaced categories. The judgment function accounts for the exponential relationship found between "ratio" estimations and "difference" ratings. An exponential judgment function explains why "ratio" judgments fit the ratio model, even though subjects are actually computing subjective differences, and it also explains why reverse attributes, such as "ratios" of easterliness and westerliness, are reciprocally related (Birnbaum, 1980, 1982). When the instructions mentioned an example "ratio" of "64," the modal and median subjects reported that the occupation of *physician* is 64 times as prestigious as the occupation of *trash collector*. However, when the largest "ratio" mentioned in the instructions was 4, the same pair received a judgment of only 4 times. Such malleability is consistent with the judgment function of subtractive theory, because in that theory, "ratio" judgments are like category ratings of subjective differences, in which the experimenter has chosen an arbitrary set of numbers for the subject to use as categories.

Thus, the results of the present study are compatible with the judgment function approach and are difficult to reconcile with either of the other approaches. If subjects are truly matching a number to a subjective ratio according either to its true ratio or to a psychophysical function for number, why are they willing to call the same "ratio" either "4" or "64"?

Can we "avoid" using examples?

To reconcile the present malleability of "ratios" with ratio theories, one might be tempted to declare that experiments such as the present one should not be done. Such a position might maintain that there is a "right way" to do psychophysics, and that experimenters should not conduct any experiment that might yield data that disagree with one or another among the several competing theories. Poulton (1979), for example, argued that contextual effects that "bias" the results should be avoided.

To save the ratio theory, some might wish to argue that the present results serve as a warning that experimenters should not mention *any* numerical examples when instructing subjects. In essence, this position asserts that any experiment that is capable of testing the ratio model should be avoided. Mellers (1983a, 1983b) and Zwislocki (1983) have debated related issues involving contextual effects due to the stimulus distribution when subjects have absolute "freedom" to use any numbers they wish. Mellers found that when subjects are given complete "freedom" to select any numbers they wish, the data show extreme variability, but still show contextual effects due to the spacing of the stimuli. Furthermore, when subjects who "freely" chose different response ranges are analyzed separately, subjects who chose to use certain response ranges behave just as subjects who are instructed or "encouraged" by examples to use those same responses. Mellers (1983b) concluded that subjects' "freedom" to choose does not add any new principle to the data, except for increasing the variability.

Birnbaum (1982) argued against the proposition that contextual effects can be "avoided." Subjects always bring some context, formed from experiences outside the laboratory, to combine with the context created by the situations the experimenter presents. To attempt to "avoid" contextual effects would be like turning off a Geiger counter at a nuclear reactor because it indicates radioactivity. Turning off the meter does not avoid the radioactivity, it merely avoids knowledge of the radioactivity. Rather than trying to "avoid" the context, it is better to systematically manipulate the context and to develop a theory to explain the effects (Parducci, 1982).

The theory that subjects compare "ratios" of the prestige of occupations by subtraction does not necessarily imply that subjects *cannot* judge ratios. When subjects are asked to judge "ratios" and "differences" of *intervals* or *distances*, there is good evidence that subjects use both of the instructed operations (Birnbaum, 1978, 1982; Birnbaum, Anderson, & Hynan, 1989; Veit, 1978). On an interval scale, ratios are not meaningful; however, ratios of intervals *are* meaningful. For example, it may not be meaningful to judge the "ratio" of the easterliness of Philadelphia to San Francisco, but it is meaningful to judge the ratio of the distance from San Francisco to Philadelphia relative to the distance from San Francisco to Denver. Indeed, evidence from the fit of two-operation theory to ratios and differences of distances leads to a scale that confirms the subtractive interpretation of simple "ratios" and "differences."

When we take results from such studies together with the present data, the evidence is consistent with the theory (a) that occupational prestige can be best represented as an interval scale in which intervals of prestige are defined but ratios are not; (b) that subjects compare prestige by subtraction; and (c) that "ratio" judgments are lawful but malleable transformations of subjective differences.

Notes

Curtis Hardin is now at Yale University, Department of Psychology, P.O. Box 11A, Yale Station, New Haven, CT 06520-7447. Correspondence concerning this article should be addressed to Michael H. Birnbaum, Department of Psychology, California State University, Fullerton, Fullerton, CA 92634. Received for publication July 13, 1988; revision received December 21, 1988.

1. Quotation marks are used for instructions to judge "ratios" and "differences" and for judgments obtained under those instructions. Quotation marks are not used for theories, nor for actual numerical computations. These distinctions are necessary because "ratio" judgments, for example, may or may not be governed by the ratio model.

2. The "ratio" responses are 100 times the example ratio; hence, a "ratio" of "4 times" corresponds to a response of 400; a "ratio" of "1/4" corresponds to 25.

3. The marginal mean for the first stimulus is given by the expression,

$$\overline{D}_{.j} = \sum_{i=1}^{r} D_{ij}/r$$

where D_{ij} equals the "difference" between stimulus s_j and s_i, and r is the number of rows. If the J function in Equation 2 is linear, then there exist constants, a and b, such that

$$D_{ij} = a(s_j - s_i) + b.$$

Therefore, $\overline{D}_{.j}$ is a linear function of s_j, as is D_{ij}; therefore, D_{ij} should be linearly related to $\overline{D}_{.j}$ with different intercepts for each level of the second stimulus, i.

4. This prediction follows because $R_{ij} = \exp(s_j - s_i)$ implies $R_{ij} = \exp(s_j)/\exp(s_i) = t_j/t_i$, where $t = \exp(s)$. Therefore, R_{ij} will be a linear function of $\overline{R}_{.j}$, because both are linear functions of t_j. However, each row of R_{ij} will have a different slope $(1/t_i)$ for each level of the second stimulus. Log transformation of "ratios' in subtractive theory is assumed to cancel the effect of the exponential judgment function, converting "ratios" to linear functions of subjective differences.

References

Attneave, F. (1962). Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 4, pp. 619–659). New York: McGraw-Hill.

Birnbaum, M. H. (1978). Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 33–74). Hillsdale, NJ: Erlbaum.

Birnbaum, M. H. (1979). Reply to Eisler: On the subtractive theory of stimulus comparison. *Perception & Psychophysics, 25*, 150–156.

Birnbaum, M. H. (1980). A comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General, 109*, 304–319.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.

Birnbaum, M. H. (in press). Scale convergence and psychophysical laws. In H-G. Geissler (Ed.), *Proceedings of the Fechner symposium*. Göttingen, F.R.G.: Hogrefe Verlag.

Birnbaum, M. H., Anderson, C. J., & Hynan, L. G. (1989). Two operations for "ratios" and "differences" of distances on the mental map. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 785–796.

Birnbaum, M. H., & Veit, C. T. (1974). Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics, 15*, 7–15.

Cross, D. (1982). On judgments of magnitude. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 73–88). Hillsdale, NJ: Erlbaum.

Dawson, W. E. (1982). On the parallel between direct ratio scaling of social opinion and of sensory magnitude. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 151–176). Hillsdale, NJ: Erlbaum.

Dawson, W. E., & Brinker, R. P. (1971). Validation of ratio scales of opinion by multimodality matching. *Perception & Psychophysics, 9*, 413–417.

DeGraaf, C., & Frijters, J. E. R. (1988). "Ratios" and "differences" in perceived sweetness intensity. *Perception & Psychophysics, 44*, 357–362.

Ekman, G., & Sjoberg, L. (1965). Scaling. *Annual Review of Psychology, 17*, 451–475.

Elmasian, R., & Birnbaum, M. H. (1984). A harmonious note on pitch: Scales of pitch derived from subtractive model of comparison agree with the musical scale. *Perception & Psychophysics, 36,* 531–537.

Gulliksen, H. (1959). Mathematical solutions for psychological problems. *American Scientist, 47,* 178–201.

Krantz, D. H., Luce, R. D., Suppes, D., & Tversky, A. (1971). *Foundations of measurement.* New York: Academic Press.

Krueger, L. E. (1989). Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences, 12,* 251–320.

Kuennapas, T., & Wikstroem, I. (1963). Measurement of occupational preferences: A comparison of scaling methods. *Perceptual and Motor Skills, 17,* 611–624.

Marks, L. E. (1974). On scales of sensation: Prolegomena to any future psychophysics that will be able to come forth as a science. *Perception & Psychophysics, 16,* 358–376.

Mellers, B. A. (1983a). Evidence against "absolute" scaling. *Perception & Psychophysics, 33,* 523–526.

Mellers, B. A. (1983b). Reply to Zwislocki's views on absolute scaling. *Perception & Psychophysics, 34,* 405–408.

Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 582–601.

Mellers, B. A., Davis, D., & Birnbaum, M. H. (1984). Weight of evidence supports one operation for "ratios" and "differences" of heaviness. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 216–230.

Miyamoto, J. M. (1983). An axiomatization of the ratio/difference representation. *Journal of Mathematical Psychology, 27,* 439–455.

Parducci, A. (1982). Category ratings: Still more contextual effects! In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 89–105). Hillsdale, NJ: Erlbaum.

Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin, 86,* 777–803.

Rule, S. J., & Curtis, D. W. (1982). Levels of sensory and judgmental processing: Strategy for the evaluation of a model. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 107–122). Hillsdale, NJ: Erlbaum.

Schneider, B. (1982). The nonmetric analysis of difference judgments in social psychophysics: Scale validity and dimensionality. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 317–337). Hillsdale, NJ: Erlbaum.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes —loudness. *American Journal of Psychology, 69,* 1–25.

Stevens, S. S. (1966a). A metric for the social consensus. *Science, 151,* 530–541.

Stevens, S. S. (1966b). On the operation known as judgment. *American Scientist, 54,* 385–401.

Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Stevens' law. *American Journal of Psychology, 86,* 3–27.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273–286.

Torgerson, W. S. (1958). *Theory and methods of scaling.* New York: Wiley.

Torgerson, W. S. (1961). Distances and ratios in psychological scaling. *Acta Psychologica, 19,* 201–205.

Treiman, D. J. (1977). *Occupational prestige in comparative perspective.* New York: Academic Press.

Veit, C. T. (1978). Ratio and subtractive processes in psychophysical judgment. *Journal of Experimental Psychology: General, 109,* 81–107.

Wegener, B. (1982a). Fitting category to magnitude scales for a dozen survey-assessed attitudes. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 379–399). Hillsdale, NJ: Erlbaum.

Wegener, B. (Ed.). (1982b). *Social attitudes and psychophysical measurement.* Hillsdale, NJ: Erlbaum.

Zwislocki, J. J. (1983). Absolute and other scales: Question of validity. *Perception & Psychophysics, 33,* 593–594.

LINKED CITATIONS

You have printed the following article:

Malleability of "Ratio" Judgments of Occupational Prestige
Curtis Hardin; Michael H. Birnbaum
The American Journal of Psychology, Vol. 103, No. 1. (Spring, 1990), pp. 1-20.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28199021%29103%3A1%3C1%3AMO%22JOO%3E2.0.CO%3B2-H

References

The Direct Estimation of Sensory Magnitudes: Loudness
S. S. Stevens
The American Journal of Psychology, Vol. 69, No. 1. (Mar., 1956), pp. 1-25.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28195603%2969%3A1%3C1%3ATDEOSM%3E2.0.CO%3B2-3

A Metric for the Social Consensus
S. S. Stevens
Science, New Series, Vol. 151, No. 3710. (Feb. 4, 1966), pp. 530-541.
Stable URL:
http://links.jstor.org/sici?sici=0036-8075%2819660204%293%3A151%3A3710%3C530%3AAMFTSC%3E2.0.CO%3B2-%23

Range Effects in Psychophysical Scaling and a Revision of Stevens' Law
Robert Teghtsoonian
The American Journal of Psychology, Vol. 86, No. 1. (Mar., 1973), pp. 3-27.
Stable URL:
http://links.jstor.org/sici?sici=0002-9556%28197303%2986%3A1%3C3%3AREIPSA%3E2.0.CO%3B2-X