Multiple anchoring of category rating scales

VIKTOR SARRIS

Frankfurt University, Frankfurt am Main, West Germany

and

ALLEN PARDUCCI

University of California, Los Angeles, California 90024

A psychophysical experiment compared the effects of two different kinds of anchoring upon category ratings of the sizes of squares: (1) single anchoring in which the same square was presented on every anchoring trial, and (2) multiple anchoring in which squares of different sizes were presented on anchoring trials. Subjects did not rate the anchors, only those squares presented on alternate trials as the series stimuli. The major finding was that the two kinds of anchoring have similar effects. As with the single anchor, the multiple anchor establishes a new endpoint for the scale of judgment. The previously demonstrated relationship of increasing and then decreasing contrast as a function of the remoteness of the single anchor (Sarris, 1967, 1976) was found also for multiple anchoring.

Our understanding of rating scales rests chiefly on psychophysical studies of contextual effects (e.g., Helson, 1964; Johnson, 1972; Parducci, 1974). The attraction of the psychophysical experiment is that the scale of judgment is so dependent upon the particular stimulus values presented in the experimental session. For example, the same sized square may be rated "very large" when presented in a series of smaller squares, but "very small" when it is the smallest square in the series. Each size is rated in comparison with the other sizes presented for judgment. Thus, contextual theories of judgment can be tested by experimentally manipulating the physical values included in the stimulus series.

The scale of judgment may also be affected by the presence of other, unjudged stimuli. For example, a square of much smaller size may be interpolated between each of the regular presentations. Even when subjects are instructed not to rate this smaller square, their ratings of the regular series are higher than when the smaller square is not presented. This phenomenon is called "anchoring": The unjudged stimulus fixes the location of the scale—just as, in the nautical analogy, an anchor fixes or maintains the location of the ship.

It is not clear how anchoring works (cf. Bieri, Atkins, Briar, Leaman, Miller, & Tripodi, 1966). A common notion is that the scale is simply broadened to include the anchoring stimulus which then defines one of the endpoints of the scale. Even when the anchor does not

itself become a new endpoint, its effects upon the actual endpoint can be assessed quantitatively using the range-frequency model for judgment (Parducci, Knobel, & Thomas, 1976). An alternative interpretation is provided by the theory of adaptation level (Helson, 1964) which asserts that the anchor pulls the center of the scale toward the anchor's own value. By either of these accounts, the effects of the anchor should be greater when it is further in value from the members of the regular series.

In psychophysical experiments using a variety of different stimulus dimensions, such as weight, pitch, and size, Sarris (1967, 1976) has demonstrated that although the anchoring effects do increase as the anchor is shifted further in value from the regular series, this increase reaches a limit beyond which the anchor becomes progressively less effective. It is as though the anchor were rejected as irrelevant to the judgmental task when it is too different in value from the stimuli being judged (cf. Brown, 1953). An algebraic model of this distance effect was presented by Sarris (1976).

In Sarris' research, the anchor was always a single stimulus value, the same anchoring value being presented on every other trial of the experimental session. A recent study (Parducci et al., 1976) used a variation of this procedure: Instead of a single anchoring value, a whole set of anchors was employed; these were always presented on the odd-numbered trials but with the selected member of the anchoring set determined by a random sequence uncorrelated with the random sequence of regular series stimuli presented on even-numbered trials. This new procedure seems more representative of the anchoring in everyday life where judgments may be affected by a variety of unjudged values, not just a single, recurring value. For example,

This work was supported in part by a grant from the Stiftung Volkswagenwerk, Hannover, to the first author. We wish to thank Chris Battelle for her help in running the experiment. Requests for reprints should be sent to Viktor Sarris, Institut für Psychologie, Johann Wolfgang Goethe-Universität, Mertonstrasse 17, 6000 Frankfurt am Main, West Germany.

how a salesman rates his own commissions may be affected by the varying commissions earned by his fellow salesmen.

The aims of the present research were to determine whether the declining effect of an anchor with extreme distance would also be found with multiple anchors, whether such an effect depended upon the range of values within the series and anchoring sets, and whether it was the extreme value of the anchoring set or its average value that determined the magnitude of the anchoring effects. A further concern was whether the anchoring effect on any particular trial depended on the value of the immediately preceding anchor or whether it depended on some feature of the whole anchoring set, either its average value or its extreme endpoint.

METHOD

Stimuli

Squares of varying size were selected from the set of 24 used in previous research (Parducci & Perret, 1971). These were projected from photographic slides, the projected widths varying from 1.3 to 35.5 cm, with each step representing a factor of 1.16 in a geometric series. The projected squares were black, centered on a white 75 × 100 cm background. Groups of 10 subjects were seated approximately 1.5 to 3 m from the screen.

Series stimuli. Except for control conditions, subjects were instructed to rate only the squares presented on alternate trials. These squares, the series stimuli, formed either a narrow-range or a broad-range set, as shown in Table 1 where the values represent the ranks of the sizes among the original 24. It should be noted that both sets consist of five values, centered at Stimulus 20 but skipping alternate steps in the case of the broad-range set.

Multiple anchors. The log-range of the anchoring sets correspond to the log-range of series stimuli, i.e., anchoring and series sets always cover the same number of stimulus steps. As shown in Table 1, the multiple-anchoring sets were centered at Stimulus 5, Stimulus 10, or Stimulus 16, with either five adjacent values (narrow range) or five alternate values (wide range).

Single anchors. Four different single anchors were used: Stimuli 1, 6, 12, and 18. The first three of these correspond to the lower endpoints of the multiple anchors for the broad-range sets; Stimulus 18 is also the lower endpoint of the narrow-range series and the second smallest stimulus in the broad-range series.

Procedure

Instructions. The subjects were each given the printed instructions which were then read aloud by the experimenter: "This is a study of how people judge sizes of different squares. A series of squares will be projected on the screen, one every five seconds. You are to make a judgment of how large or small the squares seem to you on every other trial. The purpose of presenting an extra square interspersed between the squares to be judged is to try to wipe out the memory effects of previous presentations of the regular square series to be judged. Therefore, you will make your judgments only after every other trial: the first presentation of a given stimulus pair will not be judged, only every second stimulus will be judged. I will tell you when to make and record your judgments. Make your judgments using the nine categories printed at the top of your response sheet. For example, if it seems to you that a square to be judged is 'very small' in comparison with the other, regular stimuli to be judged, record the number 2; if the square seems 'slightly larger than average,' record the number 6, and so on. For each presentation, record one of the nine possible numbers, depending on how large the square seems in comparison with all the other squares to be judged. Don't forget: the first presentation of a given stimulus pair will never be judged."

All references to unjudged presentations were removed for the control conditions. For all subjects, the response sheets had spaces for 120 responses. The nine categories printed at the top of the response sheet were: 1—very very small, 2—very small, 3—small, 4—slightly smaller than average, 5—average, 6—slightly larger than average, 7—large, 8—very large, 9—very very large.

Presentation sequence. After the instructions, the first 20 slides were presented as a preview, 10 anchors and 10 series stimuli. Responses to these preview presentations were not tabulated. There were 100 additional presentations for subjects in the experimental conditions, 50 anchors and 50 from the regular series. Control subjects received either 100 presentations of the A(1-9) condition or just 50 presentations of either the narrow- or broad-range series stimuli.

Two sequences were used for each condition, randomized within the following constraints: (a) within each successive block of either the anchor or series sets, each of the five values was presented once; (b) only one-step progressions were permitted, e.g., 3 4 6 7 5 but not 3 4 5 7 6; (c) first-order sequential dependencies between anchor and series stimuli were avoided, i.e., successive anchor and series values were uncorrelated.

The projector switched slides every 5 sec, with a switching interval of .5 sec. An entire experimental session, including instructions, preview, and 100 presentations, took approximately 25 min.

Table 1
Stimulus Values*

Method	Range	Condition	Anchor	Series Stimuli
Single Anchor	Narrow	A(1), (6), (12), or (18)	1, 6, 12, or 18	18 19 20 21 22
	Broad	A(1), (6), (12), or (18)	1, 6, 12, or 18	16 18 20 22 24
Multiple Anchor	Narrow	A(3-7)	3 4 5 6 7	18 19 20 21 22
	Narrow	A(8-12)	8 9 10 11 12	18 19 20 21 22
	Narrow	A(14-18)	14 15 16 17 18	18 19 20 21 22
	Broad	A(1-9)	1 3 5 7 9	16 18 20 22 24
	Broad	A(6-14)	6 8 10 12 14	16 18 20 22 24
	Broad	A(12-20)	12 14 16 18 20	16 18 20 22 24
Control	Narrow	Control		18 19 20 21 22
	Broad	Control		16 18 20 22 24

*Ranks in log series from 1 to 24.

Experimental design. As shown in Table 1, the experimental conditions formed two factorial designs. For multiple anchoring, the three anchoring sets were combined, respectively, with each of the two series ranges in each of two sequences (a 3 by 2 by 2 design). For the single-anchor conditions, there were four anchors (a 4 by 2 by 2 design). Controls judged either of the two series sets or else the most extreme of the broad-range multiple-anchoring conditions, A(1-9).

Subjects

Undergraduates participated as part of the regular requirement for Introductory Psychology at UCLA. Each condition used between 8 and 10 subjects, usually run in a single group. Control conditions using only the series stimuli were run with two groups to provide 20 subjects each.

RESULTS AND DISCUSSION

Overall Anchoring Effects

The major results are presented in Figure 1, where each point represents the overall mean of the ratings of those three stimuli common to all conditions, viz., Stimuli 18, 20, and 22.

Single anchors. As shown in Figure 1A, each of the single anchors elevates the ratings above those values obtained in the corresponding control condition (in which the same five series stimuli were presented without the anchor). This contrast effect is greatest when the anchor is within the regular series (Anchor 18), and it declines with increasing distance of the anchor. However, even the most distant anchor from the regular series (Anchor 1) produces significant contrast, $t(58) = 2.12$ and 3.76 for the narrow- and wide-range conditions, respectively, $p < .05$. The linear component of the decline with distance is also statistically significant, $F(1,136) = 8.04$, $p < .01$. Although ratings are higher for the narrow- than for the broad-range conditions, $F(1,136) = 5.34$, $p < .02$, range does not interact significantly with anchor distance, $F < 1$. The effects of stimulus sequence were not significant in this or in any of the other analyses.

The decrease in the power of the more distant

anchors reproduces the most novel part of the earlier distance trend with single anchors (Sarris, 1967). What is surprising is that Anchor 18, which is within the stimulus series, should have the biggest contrast effect of all. Interpretation of this finding and also of the effects of range will be presented below when we compare the ratings of different series stimuli.

Multiple anchors. Figure 1B shows that multiple anchoring also raises the ratings, as compared with the corresponding control conditions. Even the most distant of the multiple anchors, A(3-7) and A(1-9), produce significant contrast effects, $t(58) = 2.82$ and 4.18 for the narrow- and wide-range conditions, respectively, $p < .01$. However, here the distance effect is quadratic, $F(1,106) = 1.99$, $p < .05$, rather than linear, $F < 1$. This quadratic relationship is more obviously consistent with Sarris' earlier results using single anchors. As with single anchors, range is highly significant, $F(1,106) = 14.98$, $p < .001$, and is again independent of the anchoring values, $F < 1$.

Single vs. multiple anchors. All data for Figure 1, with the exception of those for Single Anchor 18 and the control conditions, were combined in an overall analysis in which the three levels of the multiple anchors were grouped, respectively, with the three levels of the single anchors. The logic of this grouping rests on the assumption that Multiple Anchor (1-9) should correspond in its effects to Single Anchor 1, each defining the same lower endpoint of the scale; similarly, Anchor (6-14) corresponds to Anchor 6, and Anchor (12-20) to Anchor 12. This correspondence is only approximate for the narrow-range conditions, but in each case the single anchor is grouped with that multiple anchor having the closest lower endpoint.

The effect of the distance of the anchor was again significant, $F(2,210) = 4.63$, $p < .02$, with most of the systematic variance in the quadratic component, $F(1,210) = 6.32$, $p < .02$. The effects of task (multiple vs. single anchor) were not significant, $F(1,210) = 1.14$, $p > .05$. Task did not interact significantly with the

Figure 1. Effects of distance of anchor upon mean rating of Stimuli 18, 20, and 22.

distance of the anchor, nor was the Task by Anchor by Range interaction significant ($F < 1$ for both tests). Thus, the overall trend is quadratic for the effect of the distance of the anchor from the series, and the difference between the effects of distance for multiple and single anchoring is not statistically significant.

Sequential Effects

The basic similarity between multiple and single anchoring is supported by the absence of first-order sequential effects for the three broad-range multiple anchors. Separate tabulations were made of ratings of the smallest series stimulus, Stimulus 18, for those trials that followed immediately after either the lowest anchoring stimulus, e.g., Stimulus 12 in A(12-20), the middle anchoring stimulus, e.g., 5 in A(1-9), or the highest anchoring stimulus, e.g., 14 in A(6-14). Although Stimulus 18 is on the average rated more than a .5 category step higher with A(6-14) than with either A(1-9) or A(12-20), differences contingent on the particular value of the immediately preceding anchor were less than a .1 category step and without systematic relationship. If each of the separate anchoring values had contributed in proportion to its own distance from the series, as assumed by adaptation-level theory, then the ratings should have been systematically higher following the smallest than following the largest among any particular set of anchoring stimuli. The simplest interpretation of the present data seems to be that only the extreme value of the anchoring set is effective and that its effect is to redefine the endpoint of the rating scale.

Rating Scales for Series Stimuli

Figure 2 presents, for each condition, the mean rating of each of the separate series stimuli. It should be noted that although Stimuli 18, 20, and 22 are common to all conditions, they span the whole range for the narrow-range conditions but only half the range for the broad-range conditions. Because the slope of the rating scale tends to be inversely proportional to range (Parducci, 1974), the abscissa has been expanded for the left half of Figure 2 to facilitate visual comparisons between ranges.

Range effects. Insofar as the smallest and largest of the available rating categories are matched to the smallest and largest of the series stimuli, slopes should be the same for the left and right halves of Figure 2. However, previous research with the same stimuli (Parducci et al., 1976) suggests that the size of the background field, which is common to both narrow- and broad-range conditions, also anchors the scale, so the same stimulus tends to elicit the same rating, regardless of range. The actual ratings reflect a compromise between the adaptation to range and absolute anchoring by the background. This is most clearly seen in the control ratings: although the control slopes are depicted in Figure 2 as steeper for the broad than for the narrow range, the difference is less than it would

Figure 2. Effects of anchors on mean rating of each stimulus in regular series. Stimulus scales equated for 18-22 and 16-24 ranges (left and right sides of figure, respectively).

have been if there were no adaptation to range. An analysis of variance performed for the control conditions on ratings of the three common stimuli demonstrates that this Range by Stimulus interaction is statistically significant, $F(2,140) = 15.10$, $p < .001$. Although Stimulus 18 is rated close to 3 in both control conditions, Stimuli 20 and 22 are rated much higher when presented in the narrow range, with the difference reaching 1.5 category steps for Stimulus 22. This interaction is inconsistent with adaptation-level theory, for it is clearly a range effect. Nor do the endpoints of the scale correspond to the extreme values of the control series. For both ranges, the most extreme of the category ratings, 1 and 9, are reserved for stimulus values 1 or 2 steps beyond the values actually presented. The net effect of the interaction between Range and Stimulus is reflected in higher overall mean ratings for the narrow-range conditions.

Anchoring effects. All four quadrants of Figure 2 show convergence toward the right: the effects of the anchors are greatest on ratings of those series stimuli closest to the distant anchor. This nonparallelism is inconsistent with adaptation-level theory (cf. Anderson, 1974, p. 275), which assumes that any effect of an anchor is upon the center of the scale, i.e., upon which

stimulus would be rated "5" in the present experiment; ratings of other stimuli are assumed to depend only on their difference from adaptation level so that the functions in Figure 2 should all be parallel. Instead, as is particularly clear in comparison with the control conditions, anchors pull the lower endpoint of the scale down, i.e., raise the ratings of the lower stimuli, but have much less effect upon the upper endpoint. This seems much more consistent with the notion that anchors redefine the stimulus range, stretching it toward their own values.

In apparent contradiction to previous findings (Sarris, 1976), it is the single anchor within the regular series, Anchor 18, that has the greatest effect. However, this may be explained by the fact that even in the 18-22 control condition, Stimulus 18 is rated about "3-small," as close to the middle of the scale ("5-average") as to the lower endpoint ("1-very small"). The effect of Anchor 18 seems to be to anchor the center of the scale so that Stimulus 18 is rated "5-average" rather than "3-small." For this special case, the effects of the single anchor appear to be more consistent with adaptation-level theory.

Apart from the extreme effects of Anchor 18, the various anchored scales are surprisingly similar to each other and well above the ratings for the unanchored, control conditions. That the upper limit for this kind of anchoring has been closely approached is suggested by the data from the additional control condition in which subjects were instructed to rate the anchor stimuli, Anchor (1-9), along with the regular series. The resulting scale was not included in Figure 2 because it is so similar to those for the regular wide-range multiple-anchor conditions in which the anchor was not rated, being intermediate between the A(6-14) and A(12-20) scales. What this means is that the contrast effects obtained with these multiple anchors are about as great as those obtained when stimuli in the most extreme anchoring set are presented as regular members of the series.

Conclusion

Taken as a whole, the present results demonstrate that the anchoring previously demonstrated using single anchors is also obtained using multiple anchors. This greatly extends the domain to which the principles of anchoring, as yet only tentatively formulated, would seem applicable. For, although the unjudged stimuli in everyday environments may sometimes be represented by a single value, comparable to a single anchor, more typically they occur over a wide range of values, like multiple anchors.

Returning to our earlier example of a salesman, it seems likely that how large or how small he rates

a particular commission would be determined in part by how it compares with other commissions he himself has been earning. His scale might also be anchored by the incomes of other salesmen: his own commissions seem larger when he is associating with those earning less than he does; however, his scale of judgment is less likely to be extended by hearing about others who work in an entirely different bracket of income. Using experimental manipulation of the context for judgment, psychophysical studies provide a basis for understanding how scales are affected by anchors at various distances from the stimuli being judged.

Within the limits of the present experiment, it seems that it is the extreme value or endpoint of the multiple-anchoring set that produces the predominant contrast effect and that this effect is greatest at some intermediate distance from those stimuli actually being rated. Besides extending the results of single anchoring to multiple anchoring, the present data give additional support to two of the empirical generalizations about anchoring: (a) anchors outside the regular series redefine the endpoints of the scale, and (b) the degree to which the scale is stretched toward the anchoring value first increases and then decreases as a function of the distance of the anchor from the old endpoint of the scale.

REFERENCES

- ANDERSON, N. H. Algebraic models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974.
- BIERI, J., ATKINS, A. L., BRIAR, S., LEAMAN, R. L., MILLER, H., & TRIPODI, T. *Clinical and social judgment: The discrimination of behavioral information*. New York: Wiley, 1966.
- BROWN, D. R. Stimulus similarity and the anchoring of subjective scales. *American Journal of Psychology*, 1953, **66**, 199-214.
- HELSON, H. *Adaptation-level theory*. New York: Harper & Row, 1964.
- JOHNSON, D. M. *A systematic introduction to the psychology of thinking*. New York: Harper & Row, 1972.
- PARDUCCI, A. Contextual effects: A range frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974.
- PARDUCCI, A., KNOBEL, S., & THOMAS, C. Independent contexts for category ratings: A range-frequency analysis. *Perception & Psychophysics*, 1976, **20**, 360-366.
- PARDUCCI, A., & PERRETT, L. R. Category rating scales: Effects of relative spacing and frequency of stimulus values. *Journal of Experimental Psychology Monograph*, 1971, **89**, 427-452.
- SARRIS, V. Adaptation-level theory: Two critical experiments on Helson's weighted-average model. *American Journal of Psychology*, 1967, **80**, 331-344.
- SARRIS, V. Effects of stimulus range and anchor value on psychophysical judgment. In H. G. Geissler & Y. M. Zabrodin (Eds.), *Advances in psychophysics*. Berlin: VEB Deutscher Verlag der Wissenschaften, 1976.

(Received for publication January 5, 1978;
revision accepted April 5, 1978.)