

# The effect on categorical ratings of personal descriptors with list length as a potential context effect

PAUL S. SIEGEL, JEFFREY ANDRULOT, and SHARON K. CALHOON  
*The University of Alabama, Tuscaloosa, Alabama*

College student subjects rated 20 personal descriptors (adjectives) on an 11-point *favorable-unfavorable (good-bad)* categorical scale. They rated using both a short and a long form. The short form presented 20 descriptors; the long form presented 83 additional descriptors for a total of 103. In two experiments, the ratings of the 20 common adjectives were found to correlate highly, suggesting the absence of a context effect. The results were viewed as extending the boundary conditions of the hypothesis developed by Hicks and his associates, which predicts no context effect when, as true of the present studies, ratings are couched in familiar or conventional terms.

That context can exert heavy influence on the judged value of various stimuli within an array has been recognized since psychologists first gave attention to the psychophysical puzzle. Quite commonly, context has been manipulated experimentally by varying some specifiable physical dimension, such as stimulus intensity. Common key stimuli are then assessed relative to two or more stimulus sets differing in average level of the dimension. A contrast effect is often found. For example, Underwood (1966, pp. 198-201) reported that the judged density of dots contained within a central square is influenced by the density of the dots contained in the surrounding area. Specifically, he found the relation to be negative; the denser the surround, the less dense was the central area judged to be. Typically, such effects have been viewed as reflecting a kind of perceptual distortion.

More recently, Hicks and his associates (Hicks, Kinbourne, & Gualtieri, 1984; Hicks, Schroeder, Gualtieri, & Mayo, 1983; Park & Hicks, 1980) have recognized another possible interpretation of context effects. They have given primary attention to the response terms required of the subject in reporting his/her judgment. In a series of studies, various stimulus sets (duration, line length, personal attractiveness, verbal expressions of amount, etc.) have been manipulated with common key stimuli judged within different overall average contextual levels. Rather consistently they have reported no context effects when judgments were expressed in familiar evaluative terms (e.g., duration expressed in seconds or line length in inches). However, when judgments were framed in cross-modal language (e.g., duration expressed as line length or line length expressed as a pie slice), contrast effects were found. Thus the higher the average duration of the context, the shorter the line length used by the sub-

ject to represent the duration of key stimuli common to three different context levels.

The first author of the present paper has for some time been concerned with a conflict paradigm that requires the subject to choose between two adjectives (*handsome* vs. *rich*, *immoral* vs. *dishonest*, etc.) that might be used by others to describe the subject (Arkoff, 1957; Murray, 1975; Schill, 1966; Siegel & McCord, 1982). In a typical procedure, long lists of personal descriptors are first rated individually on a *favorable-unfavorable (good-bad, desirable-undesirable)* categorical scale. A much briefer set is then chosen for presentation in a pair-comparisons procedure (conflict model) that permits the recording of choice times. The criteria for selecting the shorter list are the conventional psychometric considerations: the mean values must spread across the scale and present little variability. Of concern here is the fact that in moving from the long list to the briefer list, the context has been severely changed. Have the rated values of the adjectives changed? It would appear that we are dealing here with another type of possible context effect. Expressed generally, the question becomes: When a large set of personally descriptive adjectives ( $N+$  items) is reduced to a much briefer set ( $N$  items), do the *favorable-unfavorable* ratings of the crucial items ( $N$ ) change as a function of altered context?

## EXPERIMENT 1

### Method

Subjects were recruited from two sections of general psychology and were given a small course credit bonus for participation. One hundred three personal descriptors (adjectives such as *attractive* and *worthless*) were presented in a vertical arrangement within an 11-point categorical rating form. Ratings were done in class, and the subjects first read over the list of adjectives before turning a page to view the rating form. Corrections (changed ratings) were permitted ("cross out or erase").

The end points of the rating scale (-5 and +5) were labeled *Very unfavorable* and *Very favorable*. A zero was presented in the middle of the scale (-5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5), and the

The authors' mailing address is: Department of Psychology, University of Alabama, P.O. Box 2968, Tuscaloosa, AL 35487-2968.

subjects were instructed to use it to rate the word when judged to be *neither favorable nor unfavorable*. The instructions asked the subject to "imagine that people are using the word to describe other people," to decide how favorable (good) or unfavorable (bad) each word seemed to be, and then to circle the appropriate number.

Approximately 1 month later, a much briefer set of 20 adjectives was rated by most of the individuals of the same group. The instructions and format remained unchanged; only the length of the list differed from the earlier administration. These 20 adjectives were selected from the original broader set of 103 in the following procedure: The mean rating of each of the 103 words was calculated for the subjects ( $N=37$ ) completing the scale. The 20 words from this distribution that spread across the scale and presented acceptable standard deviations were selected. These adjectives are presented in Table 1, Column 1.

## Results and Discussion

Two correlational analyses were performed on the ratings yielded by subjects who completed both the earlier (long) form and the later (short) form ( $N=26$ ). In the first analysis, each word was separately analyzed. The ratings yielded on the long form were correlated with the ratings given on the short form. Table 1 presents these findings in Columns 2 and 3.

In the second analysis, the two distributions of adjectives were correlated (group data). That is, the long form mean values yielded by the group for each of the 20 words were correlated with the mean values given on the short form. The product-moment correlation was found to be .990; the rank-difference correlation was .975.

The group data correlations (.990 and .975) are remarkably high, suggesting little or no effect from the diminished context. Individual words fared worse, of course, with 6 of 20 product-moment correlations and 13 of 20 rank difference coefficients reaching the .07 level.

Complicating the interpretation of the present findings is the fact of confounding with order; the lists were rated consecutively, long followed by short. A subject's

memory of the initial rating might exert influence on the second rating, perhaps enhancing the correlation as a consistency response. Control for this is addressed in Experiment 2.

## EXPERIMENT 2

### Method

The same two rating scales (103 and 20 items on long and short forms, respectively) were used. Undergraduate subjects were recruited from an abnormal psychology class. Participation was voluntary and, again, a small course credit bonus was awarded.

One half of the subjects first rated the items on the long form; the other half rated those on the short form. Approximately 4 weeks later, conditions were reversed. Those first rating items on the long form now used the short form; those first rating items on the short form now used the long form. The design may be described as a crossover or split-plot.

### Results and Discussion

The data yielded by 19 subjects who completed both rating forms were first analyzed by ANOVA, permitting assessment of possible adjective differences, order and list length (context) effects, and, of course, their interactions.

Only the main effect of adjective differences reached significance [ $F(19,285) = 26.14, p < .001$ ]. No other main effects and no interactions were found to be statistically reliable. The interaction of adjectives and list length (context) approached significance [ $F(19,285) = 1.47, p = .095$ ]. This value fell to 1.37 ( $p = .138$ ) when what appeared to be a scale confusion error was corrected: One subject rated *harsh* +5 on the long form and -4 on the short form. The +5 was changed to -5 in the second analysis.

For each word, ratings on the short and long forms were correlated ( $N=19$ ) using both the product-moment and the rank-difference methods. Columns 4 and 5 of Table 1 present these findings. In this analysis, *harsh* was not "corrected."

Correlational analyses of the two distributions of mean values (group data) for the long and short forms yielded values of .972 (product-moment) and .929 (rank-difference).

Again, the evidence strongly supports the conclusion of no reliable effects on the ratings attributable to differing contexts: No interaction of adjective ratings with order or with list length was seen in our variance analysis. Only a main effect for adjectives was found. Furthermore, expressed as mean or group values, ratings of the 20 adjectives common to the two list lengths correlated almost perfectly (.972 and .929). Individual words also exhibited considerable consistency, with 15 of the 20 correlations reaching the .07 level for both the product-moment and the rank-difference analyses.

## GENERAL DISCUSSION

We conclude that the results obtained in these two experiments provide compelling evidence for the essential stability of adjective ratings across two contexts differing strongly in number of imbedded or contextual items. We must, of course, limit this conclusion to the categori-

Table 1  
Correlations of Common Words, Long and Short Lists

Word	Experiment 1 ( $N=26$ )		Experiment 2 ( $N=19$ )	
	Product Moment	Rank Difference	Product Moment	Rank Difference
afraid	.361*	.456*	.681*	.680*
bitter	.516*	.508*	.608*	.667*
calm	.134	.321	.354	.515*
capable	-.012	.351	.528*	.549*
careless	.206	.163	.652*	.732*
clever	.291	.323	.571*	.392
determined	.325	.398*	.543*	.488*
friendly	.137	.368*	.424	.423
frightened	.243	.279	.698*	.742*
greasy	.059	.445*	.698*	.447*
happy	.223	.456*	.034	.237
harsh	.421*	.452*	.635*	.462*
inadequate	.243	.476*	.346*	.588*
moral	.535*	.638*	.795*	.806*
nervous	.329	.342	.890*	.858*
relaxed	.123	.256	.076	.297
sad	.435*	.524*	.579*	.600*
satisfied	.326	.386*	.192	.268
willing	.230	.426*	.578*	.669*
worthless	.378*	.528*	.796*	.489*

\* $p < .07$ .

cal scaling procedure followed here. Perhaps such constancy would fail with words scaled in a pair-comparisons (or some other) methodology.

Finally, it should be noted that our subject expressed judgments (ratings) in response terms (*favorable-unfavorable, good-bad*) quite familiar to them. In this condition, alteration in the number of filler items in the rating context exercised little or no effect on the distribution of ratings. And this finding accords well with the argument developed by Hicks and his associates (Hicks et al., 1984; Hicks et al., 1983; Park & Hicks, 1980).

#### REFERENCES

- ARKOFF, A. (1957). Resolution of approach-approach and avoidance-avoidance conflicts. *Journal of Abnormal & Social Psychology*, *55*, 402-404.
- HICKS, R. E., KINBOURNE, M., & GUALTIERI, C. T. (1984). Response process in judgmental contrast. *Journal of General Psychology*, *112*, 79-87.
- HICKS, R. E., SCHROEDER, S. R., GUALTIERI, C. T., & MAYO, J. P. (1983). Effects of response language and stimulus context upon judgments of visual and temporal extent. *American Journal of Psychology*, *96*, 365-375.
- MURRAY, E. J. (1975). Resolution of complex decisional conflicts as a function of avoidance. *Journal of Research in Personality*, *9*, 177-190.
- PARK, C. P., & HICKS, R. E. (1980). Judgments of semantic stimuli as a function of context. *Journal of General Psychology*, *103*, 273-285.
- SCHILL, T. (1966). The effect and type of induced conflict on conflict generalization and later preference for conflict stimuli. *Journal of Personality*, *34*, 35-54.
- SIEGEL, P. S., & MCCORD, D. M. (1982, March). *Analysis of choice behavior in conflict situations*. Paper presented at the Southeastern Psychological Association, New Orleans, LA.
- UNDERWOOD, B. J. (1966). *Experimental psychology*. New York: Appleton-Century-Crofts.

(Manuscript received for publication October 3, 1986.)