WESTERN SPEECH

The Official Journal of
The Western Speech Communication Association

Volume XXXVI, Number 4 *Fall, 1972*

CONTENTS

Western Speech is published four times a year—Winter, Spring, Summer, and Fall—by The Western Speech Communication Association at Portland, Oregon, and printed by the University Printing Department, University of Oregon, Eugene, Oregon. Subscription, including membership in The Western Speech Communication Association, is $7.00 per year, $12.00 two years; sustaining membership, $9.50 per year, $17.00 two years; library membership, $7.00 per year, $12.00 two years; student membership, $3.00 per year; single copies, $1.50 for numbers less than ten years old and $2.50 for older numbers. Subscriptions and orders for individual copies should be addressed to Dr. Robert W. Vogelsang, Department of Speech, Portland State University, Portland, Oregon 97207.

The Intensity Component of Semantic Differential Scores for Measuring Attitude

WILLIAM E. ARNOLD, JAMES C. MCCROSKEY, AND SAMUEL V. O. PRICHARD*

THE SEMANTIC DIFFERENTIAL has achieved wide acceptance by researchers interested in measuring attitude and meaning. Several empirical studies have indicated the semantic differential is a valid and reliable attitude measuring instrument.[1]

One of the assumptions underlying the presumption of validity of the semantic differential as a measure of attitude is that the polarity of summated scores is representative of attitude intensity as well as attitude direction. This assumption has been challenged by two recent studies. Weksel and Hennes obtained correlations between semantic differential polarization scores and separate intensity scores for two samples of college students and samples of tenth- and sixth-grade students.[2] The correlations ranged from .31 to .62. Weksel and Hennes argued that these correlations indicate the semantic differential does not adequately measure attitude intensity.

Peabody drew a similar conclusion with regard to bipolar Likert-type scales and, in addition, suggested that extremeness of response is a "very general individual characteristic." Thus, differences in extremeness between persons "represent primarily response sets characteristic of the individual."[3] As suggested by Weksel and Hennes, if Peabody's conclusion were correct for bipolar Likert-type scales it would also be correct for semantic differen-

* Mr. Arnold (Ph.D., The Pennsylvania State University, 1966) is Associate Professor of Information Sciences at Illinois State University; Mr. McCroskey (Ed.D., The Pennsylvania State University, 1966) is Professor of Speech at West Virginia University; and Mr. Prichard (M.A., University of Redlands, 1949) is Instructor of Speech at The Pennsylvania State University.

[1] Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, *The Measurement of Meaning* (Urbana, 1957); Murray S. Miron, "The Influence of Instruction Modification upon Test-Retest Reliabilities of the Semantic Differential," *Educational and Psychological Measurement*, 21 (1961), 883-93; James C. McCroskey, "Experimental Studies of the Effects of Ethos and Evidence in Persuasive Communication," Diss. Pennsylvania State 1966.

[2] William Weksel and James D. Hennes, "Attitude Intensity and the Semantic Differential," *Journal of Personality and Social Psychology*, 2 (July 1965), 91-94.

[3] Dean Peabody, "Two Components in Bipolar Scales: Direction and Extremeness," *Psychological Review*, 69 (1962), 72, 73.

tials. The assumption that summated scores reflect intensity as well as direction is common to these two measures.

The conclusions of both of these studies are questionable. Weksel and Hennes base their conclusion on correlations uncorrected for attenuation. They argue that the semantic differential has consistently been demonstrated to be reliable. This is quite true. They also suggest, however, that there is little reason to assume that intensity scale responses have lower reliabilities. Quite the contrary; there is little reason to assume that the intensity scale responses do not have lower reliabilities, possibly substantially lower. Because Weksel and Hennes provide no reliability data on either of their measures, it is impossible to evaluate their obtained correlations other than to note that all of the correlations are statistically significant ($p<.01$). In light of this, the conclusion that the semantic differential does not adequately measure attitude intensity is unwarranted.

Peabody's conclusion that extremeness of response is a "very general individual characteristic" is similarly questionable. The correlations of extremeness scores obtained by Peabody were based on attitude scales designed to measure relatively homogeneous constructs. Thus, extremeness of response to the various scales should be correlated. This correlation, however, could be better interpreted as reliability of measurement than response set.

If Peabody's and Weksel and Hennes' conclusions were correct, the use of the semantic differential as a measure of attitude should be discouraged. The study reported below was designed to provide data bearing on these conclusions. Two assumptions underlie the design of the study: First, significant correlations between semantic differential extremeness scores and separate measures of intensity are presumed to indicate that summated semantic differential scores reflect attitude intensity. Second, non-significant correlations between subjects' semantic differential extremeness scores on a variety of concepts are presumed to indicate the absence of a "very general individual characteristic" (response set) of extremeness.

Thus far only the issue of attitude intensity has been considered. Involvement, or the salience of a specific issue for a given individual, should be mentioned. Studies which have controlled intensity have reported findings similar to those studies which have controlled involvement.

Mehrley and McCroskey, in a study dealing with opinionated statements and attitude intensity, demonstrated that when an individual's initial attitude position was extremely intense, opinionated statements elicited less attitude change than non-opinionated statements. Although they make no reference to involvement or social judgment theory, the results of the study correspond to predictions based on that theory.[4]

[4] R. Samuel Mehrley and James C. McCroskey, "Opinionated Statements and Attitude Intensity as Predictors of Attitude Change and Source Credibility," *Speech Monographs*, 37 (1970), 47-52.

Sereno and Mortensen and Sereno measured involvement on pretest attitude measures to determine its effect on the acceptance of discrepant communications. The first study found that highly involved subjects tended to lower their evaluations of the highly credible source more than lowly involved subjects.[5] The second study found that discrepancy and ego-involvement were interacting influences affecting human information processing.[6]

The above investigations attempted to measure either involvement or intensity, but not both. Although the present study attempted to use topics which varied in involvement, no procedure was used to measure the subjects' involvement on each issue. Thus the proposition that involvement and attitude intensity are highly related was not empirically tested in this study.

The design of the present research was influenced by procedures used by Weksel and Hennes and by Peabody even though, as we will note later, some of these procedures are questionable. However, it was felt desirable to follow procedures similar to the previous studies in order to maximize possible comparability of results. In addition to the Weksel and Hennes and Peabody procedures, the present study also employed other, more traditional, methods, which are outlined below.

METHODS

Eighty-three college students enrolled in a required basic speech course at The Pennsylvania State University were administered a series of semantic differentials using 20 concepts. The concepts were administered in two sets, each set containing 5 concepts related to campus life (high-involving) and 5 of more general interest (low-involving). The same 6 scales constituted the semantic differential for both sets of concepts. These scales were good/bad, positive/negative, fair/unfair, beneficial/harmful, right/wrong, and wise/foolish. The position and polarity of scales were randomly determined for each concept.

The first set of concepts included the following: Dormitory Life, Apartment Visitation Rule, Penn State, Cheating in College, Jammies, Ku Klux Klan, Modern Art, The Federal Income Tax, Nuclear Disarmament, and Patriotism. For this set of concepts a separate intensity measure was provided for each scale and placed adjacent to it. For example:

Modern Art

Good :____:____:____:____:____:____:____: Bad 0 1 2 3 4 5 6 7 8 9

[5] Kenneth K. Sereno, "Ego-Involvement, High Source Credibility, and Response to a Belief-Discrepant Communication," *Speech Monographs*, 35 (1968), 476-81.

[6] C. David Mortensen and Kenneth K. Sereno, "The Influence of Ego-Involvement and Discrepancy on Perceptions of Communication," *Speech Monographs*, 37 (1970), 127-34.

The Ss completed the semantic differentials for each concept according to the usual instructions and, at the same time, were told to ignore the numbers to the right of the scales. After all Ss had completed the semantic differentials for the 10 concepts, they were asked to go back to the first page of their booklets and circle the number adjacent to each scale which most nearly represented "how strongly" they felt about the way they had responded to the scale. The Ss were instructed not to change any of their previous responses to the semantic differentials, a procedure used previously by Mehling.[7]

Although the placement of the intensity measures next to the semantic differentials could have produced experimenter demand characteristics, the alternate placement of the adjectives on the semantic differentials should have mitigated against it. Further, the authors wanted to follow closely the procedures used in previous research.

The Ss' individual semantic differential scale responses were scored 3, 2, 1, 0, 1, 2, 3 to represent scale intensity. The scale intensity scores were correlated with the separate intensity scores across scales, concepts, and Ss. This procedure provided 4980 semantic differential scale intensity-separate intensity pairs for correlation.

Each Ss' semantic differential scale intensity score was also summed for each concept, as were the separate intensity scores. The summed scores for the two intensity measures were correlated for each concept and for the 10 concepts combined. This procedure provided 83 semantic differential intensity-separate intensity pairs for correlation for each concept, and 830 pairs for the 10 concepts combined. The reliability of each intensity measure was estimated by means of the Hoyt analysis of variance technique.[8]

The Ss' summed semantic differential scale intensity scores were correlated for the 10 concepts, the 5 high-salient (campus related), and the 5 low-salient (general) concepts. The same procedure was employed for the summed separate intensity scores. The correlations in the resulting matrices were transformed to Fisher Z scores.[9] The Z scores were averaged and the resulting values were transformed into the corresponding mean correlation values. This procedure provided separate mean intensity correlations for the high-salient concepts, the low-salient concepts, and the total group of 10 concepts for the semantic differential intensity measure and for the separate intensity measure.

The second set of concepts included the following: Speech 200, Draft Deferments for College Students, *Daily Collegian,* Speech 200 3-hour Research Requirement, Teenage Marriages, Capital Punishment, Freedom of Speech, Federal Control of Education, Viet Cong, and U.S. Supreme Court.

[7] Reuben Mehling, "A Simple Test for Measuring Intensity of Attitudes," *Public Opinion Quarterly,* 23 (1959), 576-78.
[8] J. P. Guilford, *Psychometric Methods,* 2nd ed. (New York, 1954), p. 383.
[9] *Ibid.,* p. 288.

A single-item intensity measure was provided for each concept in this set. This measure was a ten-step bipolar scale with one pole defined as "I am very certain of my attitude toward (concept)" and the other pole defined as "I am very uncertain what my attitude is toward (concept)." The separate intensity measure was placed below the semantic differential on the page. The Ss were instructed to complete the semantic differential and then the intensity measure.

The Ss' individual scale responses were scored as above. These scores were summed for each concept and the total scores were correlated with the separate intensity scores. This procedure provided 83 semantic differential intensity-separate intensity pairs for correlation for each concept and 830 pairs for the 10 concepts combined. The reliability of each semantic differential intensity measure was estimated by means of the Hoyt analysis of variance technique.[10]

The Ss' summed scale intensity scores were correlated for the 10 concepts, for the 5 high-salient concepts, and the 5 low-salient concepts. The same procedure was used for the separate intensity scores. The correlations in the resulting matrices were transformed to Z scores. These scores were averaged and the resulting values were transformed into the corresponding mean correlation values. This procedure provided separate mean intensity correlations for the 10 concepts, for the 5 high-salient concepts, and for the 5 low-salient concepts for the semantic differential intensity measure and the separate intensity measure.

RESULTS

In the first part of the study, the obtained correlation between the semantic differential scale intensity scores and the separate intensity scores across scales, concepts, and Ss was .42. The obtained correlations between the subjects' summed semantic differential scale intensity scores and summed separate intensity scores for the 10 concepts ranged from .35 to .63. Corrected for attenuation, the correlations ranged from .37 to .68. The obtained correlation for the 10 concepts combined was .50. Corrected for attenuation, the correlation was .55. All of the obtained correlations were significant ($p < .001$).

The obtained correlations between the Ss' summed semantic differential intensity scores for the 10 concepts ranged from -.20 to .44. The mean correlation was .12 ($p > .05$). The range for the high-salient concepts was from .04 to .26, for the low-salient concepts from -.20 to .18. The mean correlations for the high- and low-salient concepts were .15 and .05, respectively. Neither mean correlation was significant ($p > .05$).

The obtained correlations between the Ss' summed separate intensity

[10] *Ibid.*, p. 383.

scores for the 10 concepts ranged from -.04 to .49. The mean correlation was .24 ($p < .05$). The range for the high-salient concepts was from .13 to .49, for the low-salient concepts from .03 to .33. The mean correlations for the high- and low-salient concepts were .30 and .22, respectively. Both mean correlations were significant (high-salient, $p < .01$; low-salient, $p < .05$).

For the second part of the study, the obtained correlations between the 10 concepts ranged from .19 to .58. Partial correction for attenuation (no reliability estimate was available for the single item intensity measure) produced a correlation range of .22 to .60. The obtained correlation for the 10 concepts combined was .48. Partially corrected for attenuation, the correlation was .52. All but one of the obtained correlations were significant ($p < .001$).

The obtained correlations between the Ss' summed semantic differential scale intensity scores for the second 10 concepts ranged from -.10 to .46. The mean correlation was .18 ($p > .05$). The range for the high-salient concepts was from -.10 to .46, for the low-salient concepts from .12 to .38. The mean correlations for the high- and low-salient concepts were .19 and .21, respectively. Neither mean correlation was significant ($p > .05$).

The obtained correlations between the Ss' separate intensity scores for the second 10 concepts ranged from -.17 to .42. The mean correlation was .20 ($p > .05$). The range for the high-salient concepts was from .09 to .41, for the low-salient concepts from .12 to .37. The mean correlations for the high- and low-salient concepts were .28 and .23, respectively. Both mean correlations were significant ($p < .05$).

DISCUSSION

An assumption underlying this study was that significant correlations between semantic differential intensity scores and separate intensity scores would indicate that the semantic differential does, in fact, measure attitude intensity. Correlations between responses to the two measures were significant ($p < .001$) when concepts and Ss were disregarded. Correlations between the Ss' summed responses were significant ($p < .001$) when concepts were disregarded. Correlations between Ss' responses to the two measures were significant ($p < .001$) for 19 of the 20 individual concepts. Thus, the semantic differential does reflect attitude intensity.

The question, of course, is how much intensity variance can be predicted from semantic differential scores. On the surface, the above reported correlations appear to suggest that we cannot predict much intensity variance from semantic differential scores. We believe such an interpretation is unwarranted. Although the correlations reported above were corrected for attenuation whenever possible, the corrections made were quite conservative.

The reliability estimates upon which the corrections were based were internal estimates. Our experience has been that while internal reliability estimates of semantic differentials are regularly over .90, test-retest estimates after a delay of one week or more are usually nearer .80. The effect on the correlation when a less conservative correction is made is noteworthy. For example, the obtained correlation for one of the concepts in the second set was .46. A partial correction for attenuation based on the internal reliability estimate of the semantic differential intensity scores (.88) raised the correlation to .49. A test-retest reliability estimate for the semantic differential for this concept based on responses of different but comparable Ss was .81. Assuming this as the true reliability of the semantic differential, the partially corrected correlation would be raised to .51. This, of course, does not take into account the unreliability of the separate intensity measure, which in this instance was a single item measure. Such single item measures are notoriously unreliable. Even if it were as reliable as the semantic differential, however, the corrected correlation would be raised to .57. Thus, it would appear that the correlations reported above are quite conservative estimates of the relationship between semantic differential intensity scores and separate intensity measures.

It is apparent, however, that even if more liberal corrections for attenuation were generated, the correlation would not be likely to approach unity. There is substantial variance on each measure that is not predictable from the other. This may suggest to some that we should follow Weksel and Hennes' advice and administer a separate intensity measure when measuring attitude or meaning with the semantic differential.

We would question the usefulness of this procedure. The decision to use separate intensity measures requires the assumption that they are more valid measures of attitude intensity than the semantic differential. This may or may not be true. We have no empirical evidence that our intensity measures are valid. We have no empirical evidence that our intensity measures relate to involvement. Weksel and Hennes provide no such data for their intensity measures either. Independent intensity measures are not likely to be perfect measures of attitude intensity any more than the semantic differential is a perfect measure of attitude intensity. Until empirical evidence is produced which indicates that independent intensity measures correlate more closely with attitude intensity than do semantic differential scores, the conclusion that one measure of attitude intensity is substantially superior to the other is unwarranted.

Further research is needed to determine whether attitude intensity and ego involvement are independent measures of a subject's attitude toward a given topic or if they measure the same thing. We do not know whether a person can hold an intense attitude on a topic which is not involving to him.

We do know that research using involvement or intensity measures found similar results. The relationship of the two must be empirically tested.

A second presumption underlying this study was that non-significant correlations between Ss' semantic differential intensity scores on a variety of concepts would be indicative of the absence of a "very general individual characteristic" of extremeness. As reported above, obtained correlations of intensity scores were spread over a substantial range. Many of the correlations were negative. Most of the positive correlations were non-significant. Mean correlations between semantic differential intensity scores for the various groups of concepts were all non-significant, even for relatively homogeneous concepts. All but one of the mean correlations between the separate intensity scores for the various groups of concepts, however, were significant. This suggests that there is an "intensity response set" which functions in responses to independent intensity measures. This response set is not, however, reflected in semantic differential scores.

This finding is relevant to our above discussion of the relationship between semantic differential intensity scores and the separate intensity scores. It would appear that one of the reasons that the two measures are not more highly correlated is that response set is operative in responses to independent intensity measures but is not operative in responses to semantic differentials. This, of course, bears directly on the validity of the two measures of attitude intensity. The absence of an intensity response set component in semantic differential scores and the presence of that component in separate intensity measures suggest that the inclusion of additional intensity measures in conjunction with the semantic differential as a measure of attitude or meaning will most likely produce less valid rather than more valid results.

The implications of the results of this study are clear. They support the argument that semantic differential scores reflect attitude intensity as well as direction. They also clearly indicate that intensity of response to semantic differential scales is an attitude or meaning component, not a manifestation of an individual response set. The use of independent intensity measures in conjunction with semantic differential measures is unwarranted.