The Journal of Social Psychology, 1941, **13**, 123-140.

A PSYCHOMETRIC STUDY OF MEANING*

Department of Psychology, University of Florida

CHARLES I. MOSIER

The meaning of a word is a relationship between an individual, a symbol, and the object, event, or relationship symbolized. This meaning is affected, moreover, by the context of stimuli, symbolic and otherwise, in which the relationship occurs. It is supposed, in this study, that what we customarily refer to as the "meaning of a word" varies, not only from one word to another, but from one individual and from one context to another. As an example of this last, we have *"only fair," "fair work,"* and *"fair maidens."* In spite of the variation in the meaning of a given word due to the individual and to the context, there appears to remain a residuum of meaning which acts as if it were inherent in the word (not, of course, in the word *per se,* but in the accumulation of social experiences which for any word, are relatively the same from one individual and from one context to another). Thus while the word *"beautiful"* may have different meanings for different persons, and different meanings when applied to different objects, it seldom, if ever, means the same as is meant by *"ugly."* Moreover, the variations which do occur seem lawful and relatively small in magnitude.

A consideration of the origin within the individual of the meaning-relationships of a particular word indicates that these meanings result from the individual's having experienced the word in a series of situations. Thus, for a single individual and a single word, there are many meaning-relations which may be evoked. The nature of language, however, is such that, for a single word, the range of situations in which it is frequently experienced is relatively narrow, and if the meaning-relation is one that can be described quantitatively, the frequency of its appearance probably will be inversely related to the distance from the modal value. If an individual is presented with a single verbal stimulus (as in a free association test) the meaning that is evoked may be any among the repertoire of

*Received in the Editorial Office on August 10, 1939.

meanings based on his experiences with that stimulus, but it will most probably be one at, or near, the median.

The social nature of language is such that words tend to be presented to different individuals (within a homogenous language group) in approximately similar situations. It can thus be inferred that if a number of individuals are presented with a single verbal stimulus and asked to characterize its meaning in some quantitative fashion, their responses will be representative of the usual situation in which that word is used. In the light of the foregoing, it is reasonable to suppose that the distribution of their responses will be Gaussian, since the factors determining which particular meaning will be evoked are many and each small. It will be convenient to adopt this supposition as an hypothesis at a later point in the study.

In the experiment to be described, subjects were asked to rate adjectives on an 11-point scale as to their favorableness-unfavorableness. A comment here appears to be in order as to the psychological process of judging the words. The subjects did not rate the words in a vacuum, or merely as to the meaning of the word alone. Rather the rating process consisted of first, a free association test, where the adjective stimulus evokes an associated context, and second, a judgment of the meaning of the adjective in that context. This was evidenced by the comments of the subjects, and it will serve to explain some of the deviate responses. Thus for several subjects the word *"poor"* aroused the context of man, and the phrase *"poor man,"* with *"poor"* interpreted in the sense of "poverty of worldly goods" was rated as neutral or even mildly favorable. A more complete study would exercise more rigorous control of the associated context, although from the point of view of this investigation, complete variability of associated context merely serves to indicate more adequately the range within which the meaning of the word fluctuates.

We may adopt as our basic hypothesis that the meaning of the word may be considered *as if* it consisted of two parts. One component is constant for the word itself, and representative of the relatively narrow range of contexts within which the word is used; the other component is variable, and representative of variation within that range. As quantitative knowledge of the phenomena of meaning accumulates, it will be possible to enlarge the hypothesis to render it more flexible, and more nearly a picture of what seems

to occur in the process of an individual's assigning meaning to a word or a word-phrase. For the present, we are investigating the tenability of the hypothesis for determining for each word in a set of words the central tendency of its range of possible meanings, and the extent of that range.

We may quantify our hypothesis by saying that, if M be the meaning of the word for a given individual, if x be that part of the meaning which may be taken as constant from one person and one context to another, i be the part which varies from one individual to another and c be the portion of the meaning which is due to the context, then

$$M = x + i + c.$$

Furthermore, $x_{i\text{-}1} = x_{i\text{-}2}$ and $x_{c\text{-}1} = x_{c\text{-}2}$, that is, the x's are the same, regardless of the individual interpreting the word, and regardless of the context in which the word is interpreted. Finally, for a given word, any one of the three components may be zero, and for certain words, at least, x may be multiple-valued. In the present study the components i and c cannot be distinguished and may be combined and designated v (for variable). The hypothesis then becomes

$$M = x + v.$$

For the present study, words expressing judgmental relationships which could be placed along a favorable-neutral-unfavorable continuum were selected for study. Two hundred and fifty-six such words were selected from Thorndike's *Word-List* (5) and presented to subjects for their judgments as to the degree of favorableness or unfavorableness.

The *Word-List* was searched for all words that unequivocally expressed judgments of favorableness or unfavorableness, regardless of the objects to which the word might be considered appropriate. Words which were utterly strange, unusual, or which expressed judgments in only one of several widely divergent meanings were eliminated. An attempt was made to secure all of the words which were commonly used to indicate favor or disfavor. There were 256 different words selected and, by the time certain additions and duplications to be described later were made, 296 stimuli presented to the subjects for scaling. In order to test the reliability of the

method, and to obtain some indication of the effect of neighboring stimuli on the judgment of a particular word, certain words were repeated at widely different places in the list. The words repeated were *neutral, normal, excellent, desirable, disgusting, unsatisfactory.* A second addition was made in an attempt to study quantitatively the effect of the usual adverbial intensives, *quite, very,* and the like. A set of five words was selected, and these words repeated with each of the seven intensives studied to observe the effect of the intensive on the scaled meaning of the word. The words used were *desirable, agreeable, indifferent, poor,* and *unsatisfactory.* These were paired with each of the intensives *very, completely, quite, unusually, very very, extremely,* and *highly.*

These 296 stimuli were mimeographed, 45 words on a page, making seven separate pages. The position of the word in the page was fixed, but to control the order of presentation as much as possible, the pages were permuted in cyclic order in the assembling of the pages into booklets. Thus Page 1 was the first page for 1/7 of the subjects, and the last page for another 1/7 of the subjects. The cover page of each booklet was devoted to instructions. These are repeated here for the light they throw on the problem and the method.

WORD STUDY

Here are a number of words often used to express your judgment about a person, an object, or an event. These judgments range from extremely favorable to extremely unfavorable.

Opposite each word are the numbers, 1, 2, 3, . . . 9, 10, 11, and 0. Consider each word and, for those words which express the most extremely unfavorable judgments, *circle* the 1; for those expressing a judgment not quite so extremely unfavorable, circle the 2, and so on. For those words which express a judgment of indifference, neither favorable nor unfavorable, circle the 6. For those words which are mildly favorable, mark the 7, and so on. Those words expressing the most extreme of favorable judgments will be marked 11. If it should happen that you find a word which cannot, for some reason, be rated, circle the 0.

While these words are not all applicable to the same object, they do have in common the quality of favorableness or unfavorableness of the judgment expressed by the word. Thus, while "hateful" might not be applied to the same object as

would "repulsive," both are alike in expressing unfavorable judgments. It is this *favorableness or unfavorableness of judgment* which you are asked to make the basis of your rating.

Remember: 1 means most unfavorable
6 means neither favorable nor unfavorable
11 means most favorable
0 means cannot be rated.

Try to keep the steps between 1 and 6, 6 and 11 equal as far as differences of favorableness are concerned.

Work carefully, but do not study each word too long. Record your first impression.

These written instructions were supplemented by verbal instructions, found necessary after a few trial administrations: that the basis of the rating should be the meaning the subject understood as he read the word in a book, rather than the meaning which he might intend as he used the word; and that words with which the subject was too unfamiliar, or words which, in his opinion, did not express a judgment that could be placed on a favorableness-unfavorableness scale were to be marked 0.

The first stimulus is presented as it appeared in the booklet as a sample.

Sample Stimulus

menacing	1	2	3	4	5	6	7	8	9	10	11	0

The subjects were students in introductory and second courses in psychology. They gave evidence of being interested in the study, and made a serious effort to comply with the directions to the best of their ability. As an added precaution, the papers were studied for clear evidence of failure to comprehend the task in either of two ways. First the distribution of responses among the categories was checked for each paper, and papers were eliminated if there were fewer than three responses in any category. The validity of any results from the data depends on the symbols 1, 2, . . . 11 having approximately the same meanings for every subject. It is obvious that if one subject marks 30 of the words as "1" while another subject marks no words "1" because, as one subject expressed it *"The most unfavorable words he could think of didn't occur in the list,"* these numerical symbols or responses do not have the same significance for the two subjects. For these reasons, then, it was felt that the

validity would be increased, rather than decreased, by rejecting the results of those subjects who consistently failed to mark responses 1 or 11, or who consistently marked all words either 1, 11, or 6.

Second, the papers were checked for evidence of failure to understand the basis on which the words were rated, and papers were rejected if the subject returned more than five obviously absurd responses, e.g., "extremely unsatisfactory" marked 11, or "excellent" marked 2. Allowance was made for mistakes in markings, possible ambiguities in the interpretation of the word, and ignorance of the word. Only those papers were eliminated in which there was unmistakable evidence that the subject was responding on some other basis than the degree of favorableness or unfavorableness of the judgment expressed by the word. The bases for ratings which the rejected subjects later reported as governing their responses ranged from their personal liking or disliking for the word, to an admission of horseplay, and random marking on no basis whatever.

For each word there were approximately 140 judges (the number varied with the number of 0's marked, complete omissions, and incomplete papers). Tests of stability of scale values from successive samples of 50 cases indicated that with 150 cases, the median value of a word was stable within a range of .2 of an interval.

In the course of the study, one interesting fact became apparent, at first from the comments of the subjects, later from an examination of the papers for rejection of those in which the directions were grossly misunderstood. This was consistent evidence of individual differences in the interpretation of the extent of the scale. This may be merely a difference in the meaning attached to the numerical symbols representing judgment, or it may indicate personality differences describable as extremism and conservativism. A check on the existence of the tendency is obtainable from a correlation between the ratings of two words. From these correlations, both for words adjacent in the stimulus sheets, and for words widely separated, confirmation of the hypothesis was obtained. Words with median values on the same side of the neutral point, e.g., two favorable words, correlated to the extent of a coefficient of .30-.40, indicating that individuals who marked the first word more favorably than the average of the group, tended also to mark the second word more favorably than did the group. For two words with median

values on opposite sides of the neutral point, the correlation coefficients were of the same magnitude, but of negative sign, indicating that the individual who marked the first word more favorably than did the group, tended to mark the second word more unfavorably. The correlation coefficients indicate nothing of course, beyond the fact that for one individual, there was a tendency to mark words either at the extremes, or toward the neutral value. It seems probable that a more careful definition of the meaning to be attributed to each numerical symbol,[1] 1 . . . 11, would reduce this tendency and materially reduce the recorded ambiguity of the words. It opens the possibility that the magnitude of ambiguity recorded for a given word is not entirely due to the ambiguity of the word, but to a considerable extent to the ambiguity (for the total group of subjects, though probably not for the individual subject) of the symbols used to convey the judgment. Such an effect would not, in all probability, affect the *relative* ambiguities of the stimuli.

When the data were complete, frequency distributions of the responses to each word were prepared and the method of equal appearing intervals applied to the scaling of each word. This method yields a scale value and a measure of the ambiguity of the word, but it is subject to the rigorous assumption, which does not seem warranted for these data, that the steps 1-2, 2-3 . . . 5-6, 6-7 . . . 9-10, 10-11 are psychologically equal. It is subject, moreover, to the well-known "end-effect," whereby the scale values of words near the ends of the scale are pushed toward the neutral values, and the ambiguity of the word is made artificially small. Since the much more precise method of successive intervals was also found applicable, discussion of the results of scaling will be in terms of the values found by that method, and will be deferred until certain phenomena apparent from a study of the distributions themselves have been considered.

The results for the words which were repeated at different places in the study indicate the reliability of the method, and, incidentally, the effect of immediately preceding stimuli on the scaling of the stimulus. The words repeated, with the frequency distributions,

---

[1]In future studies of this sort, this definition of the symbolic responses might be achieved by including in the instructions examples of the sort of stimuli to be placed in each category of judgment, based on the results of this study.

medians and quartile deviations from the two presentations are given in Table 1. The correlation coefficients between the two distributions are also given.

TABLE 1
DISTRIBUTIONS OF RESPONSES TO REPEATED STIMULI

Stimulus	1	2	3	4	5	6	7	8	9	10	11	Md.	Q.	*r*
Disgusting	43	51	32	16	3							1.6	1.6	.97
	44	43	38	17	6							1.7	1.8	
Unsatisfactory	14	39	47	31	17							2.5	1.8	.90
	9	34	36	43	26	1	1					2.9	1.8	
Neutral				1	6	133	5		1			5.5	.6	.99
			1	1	6	131	6	2	1			5.5	.6	
Normal	1			1		104	15	5	12	7	1	5.7	.9	.99
	1				2	104	19	5	10	6	2	5.7	.9	
Desirable						3	9	29	57	43	8	8.6	1.6	.90
						3	12	34	45	46	8	8.7	1.4	
Excellent							2	3	14	33	96	10.2	.9	.99
							1	3	14	42	88	10.2	.9	

These distributions will also serve as representatives of the majority of those obtained. The exceptions fell into three types which will be discussed separately.

The first type of divergence from expected results we may call ignorance of the meaning of the word—at least of its favorable-unfavorable meaning. This result is manifested primarily by a high incidence of the response "0"—"unable to rate." Many of the subjects seem to have been confused in the directions and in the case of an inability to rate the word, marked, not "0," but "6." If we consider nine as the maximum allowable number of zero responses to a single word, eliminating words with 10 or more such responses, there are 63 words eliminated; in only 26 were there 20 or more zero responses. Table 2 lists, in order of the number of omissions, the words omitted by 20 or more subjects. Inspection of the table

TABLE 2

WORDS MARKED "UNABLE TO RATE" BY 20 OR MORE SUBJECTS

1.	Propitious	8.	Perverse	15.	Ominous	21.	Peerless
2.	Cloying	9.	Ecstatic	16.	Bonny	22.	Persuasive
3.	Iniquitous	10.	Noxious	17.	Inflaming	23.	Abhorred
4.	Pernicious	11.	Solacing	18.	Calamitous	24.	Debased
5.	Expedient	12.	Seemly	19.	Odious	25.	Adverse
6.	Satiating	13.	Estimable	20.	Despicable	26.	Superlative
7.	Pestilential	14.	Seductive				

confirms the conclusion that a high proportion of omissions indicates ignorance of the word.

Since not all of the subjects who were ignorant forbore to rate the word (13 subjects marked "propitious" on the unfavorable side of neutral) and since many used "6" to mean "cannot be rated" (in spite of clear written and oral instructions) the data for these words are unreliable, and any results based on them must be accepted with caution.

The second sort of atypical results occurred in the case of words with two distinct meanings—more precisely, words usable in two widely divergent contexts. These words were marked by decided bimodality of the frequency distributions. These results do not detract from the method and are not unexpected. Analysis reveals that we are here dealing with two words which happen to be spelled and pronounced the same. In the terms of the hypothesis advanced, x, the constant component of the meaning, is multiple-valued. In terms of the analysis of the act of judging, the visual appearance of the word in one subject aroused "word-in-context-one," in which one meaning of the word was apparent; for another subject the identical stimulus aroused "word-in-context-two," in which a markedly different meaning was obvious. There were 21 such words, in addition to seven contained in the list discarded because of ignorance. As an example of such words we may cite "completely indifferent" with modes of 48 responses at Category 6 and of 37 responses at Category 1. The complete distribution is shown in Table 3. The confusion seems to have been between active and passive objects—a "completely indifferent" person is highly unfavorable, whereas an object that is completely indifferent is quite neutral. (Of the 28 bimodal words and phrases, eight were adverbial varians of "indifferent"—originally chosen as the neutral stimulus.) Every one

TABLE 3

SAMPLE BIMODAL DISTRIBUTION OF JUDGMENTS
STIMULUS: *"Completely Indifferent"*

Categories	1	2	3	4	5	6	7	8	9	10	11
Frequencies	37	20	17	12	13	48	1	0	1	0	0

of the eight variants showed a mode at 6 and a second mode at one of the unfavorable categories. *"Bearable,"* with modes at 5 and 7, *"bewitching,"* with modes at 10, 6, and 2, and *"seductive,"* with modes at 9, 6, and 4, were the only instances in which modes lay on opposite sides of the indifference point. The interpretation of the results for these words is clear. For a number of words, the bimodality is more apparent than real—an artifact due to variations in the sizes of the subjective categories (1). When the assumption of equal intervals between categories is replaced by more precise methods, the apparent bimodality disappears for seven of the words. Table 4 gives a complete list of words. It should be repeated that

TABLE 4

WORDS EXHIBITING MARKED BIMODALITY OF RESPONSE

1. Acceptable	10. Completely indifferent	19. Irresistible
2. Amazing	11. Extremely indifferent	20. Normal†
3. Appalling*	12. Highly indifferent	21. Peerless
4. Base*	13. Quite indifferent	22. Satiating*
5. Bearable	14. Unusually indifferent	23. Seductive*
6. Bewitching	15. Very indifferent	24. Sublime
7. Choice*	16. Very, very indifferent	25. Tempting
8. Important	17. Inflaming*	26. Unfit
9. Indifferent	18. Indispensable	27. Unspeakable

*These words were marked "unable to rate" by 10 or more subjects.
†This word appeared in two places in the study, exhibiting bimodality both times.

these results, far from discrediting either the method or the hypothesis, indicate that the method is adequate to detect contextual shifts in the meaning of words. The relatively small number of words which show marked abnormality indicate that for most words there is a relatively narrow range of favorableness or unfavorableness of contexts in which that word is likely to appear. The results further indicate that the distribution of the frequency of appearance of these contexts, under the conditions of a free association test, is reasonably Gaussian.

The third, and possibly the most significant sort of deviation from regularity in the results appears in a failure of the distributions to cross the neutral point. The distributions appear to give a "precipice" effect. An example will make clear what is meant. The frequencies for the word "unnecessary" are shown:

Stimulus: *unnecessary*

Categories	1	2	3	4	5	6	7	8	9	10	11
Frequencies	5	8	24	26	45	29	—	—	—	—	—

The drop from 29 to 0 is what is referred to as a "precipice." The psychophysicist will recognize here an end-effect in the middle of the scale. There are at least two possibilities to explain this phenomenon, which was nearly universal for words whose distributions lay in the middle of the range; practically no word whose distribution crossed Category 6 failed to exhibit this precipice effect. Guilford (1) has shown that the subjective size of the categories of judgment varies considerably from one to another, and it is conceivable that the neutral category was conceived by the subjects as covering a wide range in the middle of the favorable-unfavorable continuum. If this be the explanation, then the phenomenon is of little interest, since it is merely an artifact resulting from a peculiarity of the scale of judgment. Another possibility, much more interesting, though as yet without convincing evidence, is that we are dealing with a real "end-effect"—that favorableness-neutral is one continuum and neutral-unfavorable is another, not collinear with the first. (At present the writer has nothing to offer but the suggestion and the "precipice" effect to support it.)

A third possibility is that the non-linearity exists, but only as a result of the instructions. Thus it may well be that if the instructions had been in terms of "degree of favorableness," rather than in terms of "favorableness-unfavorableness," then the piling up at Category 6 and the failure of any of the distributions to cross the neutral point would not have appeared.

Certain other results, of no special significance other than their general interest, are apparent from a study of the distributions. The three words *"pitiable," "piteous,"* and *"pitiful,"* were scaled 7, 8, 9, or 10 by 15 subjects, the same ones in every case. Here is probably an instance where the effect of individual meaning, as apart from meaning due to context, is apparent.

Several words were obviously confused with other words in the minds of the judges. *"Prudent,"* for example, was judged below 6 by 30 subjects—probably due to a confusion with *"prudish."*

One phenomenon which I hesitate to describe lest it cast doubt on the validity of the entire study in the minds of the grammarians is this: The four words, *"good," "better," "bad,"* and *"worse"* were included among the stimuli to be judged. The subjects unequivocally rated *"good"* better than *"better"* and *"bad"* as worse than *"worse."* The magnitude of the differences and the consistency of its occurrence renders it certain that it was not a chance phenomenon. Startling as this may appear to a grammarian, it is psychologically sound, since *"good"* is a positive assertion, whereas *"better"* implies comparison with some standard which might, in many cases, be itself unfavorable. Compare the often heard comment, *"He's getting better, but he is still far from good."*

The word *"fair,"* used in rating scales and in reporting grades to denote a value slightly below neutral, was rated below neutral by only 7 per cent, and neutral or below by only 32 per cent of the judges. The median value was a full point above the neutral point of 5.5. For these students the contextual meaning of *"fair"* is distinctly favorable.

"Ordinary" showed a marked skewing at the low extreme, indicative of the common use of *"ordinary"* to imply inferiority. The word *"normal"* was marked 6 by 70 per cent of the group, but there was a pronounced mode at 9. That a number of the subjects were students in a class in abnormal psychology may have contributed to this result.

The word *"devilish"* was scaled as unfavorable by the vast majority of the subjects, but a small group seemingly took the word in a playful context and marked it from 7 to 10. A few subjects reacted similarly to *"maddening,"* although the favorable judgments were fewer than in the case of *"devilish."*

"Fit," although judged above 6 by the majority, was assigned every rating from 1 through 11. Since 13 subjects omitted it, it is probable that the subjects failed to understand the word as "appropriate" or "healthy," but rather confused it with the noun. While these are the most outstanding deviations from expectation, other instances of the effect of ambiguity of context might be cited.

The hypothesis thus far made—that the meaning of a word may be treated as if it consisted of two components, one fixed, and one varying with the individual and with the associated context, with the further provision that the variability of the contextual portion of the meaning is slight, relative to the range of the fixed component —is thus seen to be verified. However, the analysis previously made of the act of judging makes reasonable the following, stronger hypothesis: *That the responses to a word can be made to project a Gaussian (normal) distribution along a scale, and that the responses to all words will be normal on the same scale;* i.e., that a scale can be found such that the distributions of responses to all the words will be in accordance with the normal law of error. We may interpret this as the "random-ness" with which any particular contextual meaning will arise. The strength of the hypothesis lies in the finding of a *single* scale such that *all,* or nearly all of the words studied will yield a strictly Gaussian distribution.

The discovery of the scale and the verification of this hypothesis was carried out by a modification of Thurstone's method of successive intervals (2, 4). In this method the mean response to a standard stimulus is taken as the origin and the standard deviation of that word is taken as the unit of measure. The standard stimulus chosen was the word *"completely unsatisfactory,"* with its scale value defined as zero and its ambiguity defined as one. From eight to 14 independent determinations of the scale value of the upper limit of each category were made and the means taken as the values of those upper limits. The hypothesis may then be verified for each word by plotting the percentile values of each interval on a graph whose ordinate is scaled in percentiles spaced as on standard probability paper and whose abscissa is scaled proportionately to the scale values of the intervals. If the plot for each word is linear, then the hypothesis is verified. If the abscissae are also marked in units of the scale, the scale value and ambiguity of each word may be read directly from the graphic plot (2). Typical plots are shown in Figure 1.

Of the 296 stimuli studied, only 78 exhibited the slightest evidence of non-linearity. For the remaining 218 stimuli, then, the hypothesis that the normal law of probability will describe the frequency with which a particular meaning is evoked out of the repertoire of mean-

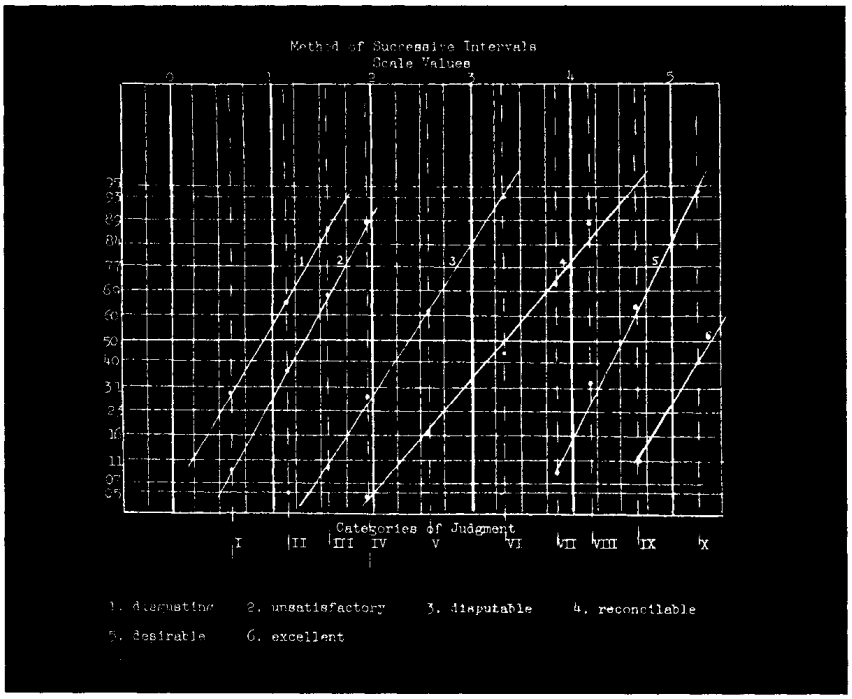

FIGURE 1

ings available to the individual (where this repertoire represents the sum total of situations and meaning-relations in which this word has been experienced by the group of subjects) is convincingly verified.

Of the 78 words which were even slightly non-linear, 24 were so nearly linear that scale values and dispersions could be computed without distortion. For these words, the non-linearity was apparent only at the extremes of the range where the proportions which constitute the observations are statistically the most unreliable. The non-linearity of the remaining 54 words was attributable to ignorance of the meaning of the word in 25 cases, to bimodality of meaning in 19 cases. There are then only 10 cases non-linear for some cause other than these two. If the precipice effect is evidence that favorableness-unfavorableness cannot be considered as extremes of the same continuum, then the non-linearity of four more words is ac-

counted for, since these words all exhibited a marked piling-up of judgments at Category 6. The words exhibiting non-linearity not due to ignorance or to bimodality of meaning are listed in Table 5. Those exhibiting the precipice effect are starred.

TABLE 5

WORDS EXHIBITING NON-LINEARITY NOT DUE TO IGNORANCE OR BIMODALITY

Alarming*	Luring*	Piteous
Exasperating	Medium*	Pitiable
Gladdening*	Ordinary	Tragic
Invigorating		

*These stimuli exhibited the "precipice" effect.

When the 218 words for which the hypothesis was verified had been scaled, there resulted a scale extending from a value of 0.00 for the words, *"completely unsatisfactory,"* and *"vile,"* to a value of 5.66 for the word *"excellent."* Table 6 presents words selected at

TABLE 6

Stimulus	Scale Value
Completely unsatisfactory	0.00
Repulsive	0.50
Disgraceful	1.00
Wrong	1.50
Unnecessary	2.00
Disputable	2.36*
Excusable	2.85*
Average	3.08
Pardonable	3.48
Comfortable	4.04
Desirable	4.50
Highly agreeable	5.02
Divine	5.50
Very, very desirable	5.66

*There were no stimuli with scale values between 2.36 and 2.85.

intervals from the entire range, together with the scale value of each word. Complete tables of all results, frequency distributions and scale values, are available (3).

The ambiguity values ranged, for the scalable words, from 1.13 for the word *"divine"* to 0.24 for the word *"neutral."* The distribution of ambiguity values had a median of 0.60, a lower quartile of 0.54 and an upper quartile of 0.74. The ambiguity values were, in

general, largest for words at the extremes of the scale, as is commonly found with other types of stimuli. These data exhibited a peculiarity, consistent with the hypothesis that favorable-unfavorable is not a single continuum, in that the ambiguity values were smallest, not at Category 6 (where they were nearly as great as at Categories 1 and 11) but at Categories 3 and 9. Again we find that the middle of the scale exhibits properties characteristic of the extreme of a scale of judgment.

The results of the study of the effect of adverbial intensives applied to adjectives are presented in Table 7. Four of the five adjectives selected are ranged across the top of the table each heading a column. (The fifth adjective, *"indifferent,"* behaved atypically because of ambiguous associated contexts.) Each row of the table presents the scale values for one of the adverbial modifiers studied.

TABLE 7

SCALE VALUES AS AFFECTED BY ADVERBIAL MODIFIERS

Modifier	Desirable	Agreeable	Poor	Unsatisfactory
(Unmodified)	4.50	4.19	1.60	1.47
Quite	4.76	4.45	1.30	1.00
Very	4.96	4.82	1.18	0.75
Unusually	5.23	4.86	0.95	0.75
Completely	5.38	4.96	0.92	0.00
Highly	5.15	5.02	*	0.71
Extremely	5.42	5.10	0.95	0.10
Very, very	5.66	5.34	0.55	0.25

*The phrase, *"highly poor,"* was not included in the list of stimuli since it seemed contrary to good usage.

Certain conclusions seem to be justified on the basis of these results. While several of these conclusions are obvious without the necessity of a study such as this, they serve to reinforce the validity of the method.—If a method yields no results consistent with previous knowledge, we rightly doubt the method rather than the previous conclusions; if, on the other hand, a study yields only results which could have been known in advance, it qualifies merely as "busy-work."—The most obvious of the conclusions is that the effect of an intensive is to shift the meaning away from the neutral point toward the extreme. The magnitude of the effect seemingly varies both with the adjective and with the modifier, but certain tendencies

can be noticed. For the adjectives studied, *"quite"* is the weakest intensive, and *"very"* is next in every case. *"Completely," "very very,"* and *"extremely"* are uniformly strong intensives, while *"quite," "very,"* and *"unusually"* are uniformly weak. *"Highly"* belongs with the weak or the strong group, depending on the adjective it modifies. The effect of a strong intensive is quite large. The shift in the scale value of an adjective produced by one of the three strong intensives is about one-fifth of the difference in meaning between *"vile"* and *"divine."*

The effect of the modifier on the ambiguity of the word is more difficult to determine, since it is obscured by the effect of scale value. There appears to be a tendency for the ambiguity of modified words to be less than that of words of similar scale values.

Summary and Conclusions

The application of the quantitative methods of psychophysics to the problem of meaning has resulted in the general confirmation of previous knowledge and in the discovery of other, not so readily predictable phenomena. The two basic hypotheses resulting from the analysis of the process of the arousal of meaning are confirmed: first, that the meaning of a word may be considered *as if* it consisted of two parts, one constant and representative of the usual meaning of the word, and one variable, representative of individual interpretation in usage and associated context and general usage; second, that the frequency with which any particular meaning is evoked is describable by the Gaussian Law.

In addition to these primary results, certain others of interest have accrued. The presence of words with two discrete meanings, yielding definitely bimodal frequency distributions of responses, has been noted. The possibility that "favorableness-unfavorableness" is not a single psychological scale is indicated. Some additional knowledge of the process of understanding and using words has been gained from a consideration of atypical and unexpected results for particular words. The effect of adverbial modifiers on the meaning of an adjective has been studied and certain generalizations warranted. A scale with a rational basis has been developed and values describing quantitatively the modal meaning and the ambiguity of more than 200 adjectives have been obtained.

Finally, the successful application of quantitative methods to a new field has opened up new problems and indicated new techniques for solving already existing problems. Adjectival characterizations of rating scales may now be more nearly spaced equally, and the definition of categories of judgment and rating may be made more precise by use of the scale values here obtained. A technique for scaling attitudes appears latent in the introduction of the attitude-object, e.g., nationalities or personalities, into a list of words to be scaled as these were. Studies in the effect of variation of context, particularly in the noun described, on the scale value of the adjective should throw further light on the phenomena and processes involved in meaning. Studies similar to this in varying language or dialect regions should reveal subtle shifts in word usage which might otherwise escape the philologist. The effect of "overworking" a particular word, as in the language fads which sweep the country, can be directly studied. Temporal changes in the meaning of a word may be of interest in the study of language growth and change.

The variations in meaning due to the individual, as exemplified in the responses to *"pitiable," "piteous,"* and *"pitiful"* might be made the basis of further study by the individual psychologist and provide a useful addition to the free-association test and the personality inventory techniques.

In general, this study indicates that rational quantitative methods can be applied to problems of meaning and language, both in the formulation and verification of hypotheses, and in the development of measurements for their practical and technical usefulness in other problems.

REFERENCES

1. GUILFORD, J. P., & JORGENSEN, A. P. Some constant errors in ratings. *J. Exper. Psychol.*, **22**, 1938, 43-57.
2. MOSIER, C. I. A modification of the method of successive intervals. *Psychometrika*, 1940, **5**, 101-107.
3. ———. Tables from a Quantitative Study of Meaning. (Privately issued.)
4. SAFFIR, M. A. A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 1937, **2**, 179-191.
5. THORNDIKE, E. L. Teacher's Word Book. New York: Teachers College, Columbia University, 1932.

Department of Psychology
University of Florida
Gainesville, Florida