SCIENCE

Vol. 104, No. 2707 Friday, 15 November 1946

Semantic Choice and Personality Structure

Florence L. Goodenough
University of Minnesota

MAN IS A RECEPTIVE CREATURE. Not only is his body equipped at every point with sensitive nerve endings by means of which physical events are transformed into nerve action, but he is further provided with a muscular system that enables him to bring himself into active contact with a wide variety of objects and events. Thus, in part through the inevitable conditions of living and in part as a result of his own purposeful activity, he is constantly bombarded by an unending series of external forces which impinge to a greater or less degree upon his sensitive organism.

He is also a retentive creature. Like footprints in moist clay, the experiences through which he passes in the course of his life leave their impression upon the structure of his personality. But because man is, first of all, an active creature—a living organism for whom contact with the environment means a constant sequence of metabolic interchange, both physiological and psychological—such an analogy cannot be pushed too far. Certainly the form and pattern of man's personality is not merely the passive result of the environmental strains and stresses through which he has passed. Not only does man leave his own footprints on the sands of time, but the imprints which the external world makes upon him are continually being subjected to internal pressures and strains arising from the changes incident to his own growth and development and to the accumulation of new experiences which overlie and modify the traces left by previous events.

It is this active and dynamic interchange between the self and its surroundings which lends to an individual's impressions of former events their chief importance as indicators of personality characteristics. If external happenings were merely to leave an exact and factual record of their overt characteristics upon the mind of the person who experiences them, a study of such records would have little more than historical significance. It is just because mental impressions are not simple reproductions of the events which gave rise to them but, on the contrary, have been restructured, reorganized, elaborated, or simplified in conformity with the character of the experiencing individual that their later contour becomes worthy of serious investigation by the psychologist. Within limits, at least, the psychological interest in such impressions increases in proportion as their factual accuracy diminishes.

Since the days of Jung it has been recognized that the nature of these reconstructed memory traces may be explored through the medium of free associations to common words. The use of word association tests in clinical practice was given a certain degree of statistical precision by means of the standardized frequency tables compiled by Kent and Rosanoff (*4*), who demonstrated that mentally disturbed persons who are finding the real world so difficult or unsatisfying that they attempt to withdraw into a universe of their own contriving are likely to give an excess number of responses not included in the standard lists but suggested by their own individual problems and fantasies. Many other diagnostic signs were noted by these authors and others.

Most of the studies of word association carried out up to the present time have aimed chiefly at devising methods for clinical diagnosis. Behavioral manifestations of one kind or another are noted in connection with responses to particular stimulus words, and these leads are followed up in an attempt to uncover the meaning of such deviant responses in terms of the specific associations which gave rise to them.

That studies of this kind have been, and will continue to be, of decided clinical usefulness there can be no doubt, nor can the fact be questioned that important scientific information has been gained in this way. But there is another aspect of the matter which thus far has been much less thoroughly explored. It is reasonable to suppose that the restructuring of remembered experiences by the individual is not

Vice-presidential address delivered at St. Louis, 28 March 1946, before the Midwestern Psychological Association, meeting in collaboration with Sections I and Q, AAAS.

merely a fortuitous activity, but the resultant forms and patterns are in large part determined by the subject's individual personality. It seems altogether likely, therefore, that not merely those responses which are so bizarre or unusual as to attract the attention of the experimenter but also the particular pattern of associations which fall within the normal or usual range may have diagnostic significance with reference to the broad aspects of personality which have resulted from the continuous assimilation of new experiences by a growing and changing organism.

This idea is, of course, not new. As early as 1925 Wyman (*11*) made use of the free association method in the development of a group test for school children which was designed to measure three kinds of interest, designated by her as "intellectual interest," "social interest," and "activity interest." More recently, Terman and Miles included a controlled association test in the multiple choice form as a part of their test for the measurement of mental masculinity and femininity. Maller (*6*) has constructed a test of the same type but using only two alternatives for the measurement of neurotic tendencies. A number of other investigators have experimented with both free and controlled association as a device for the appraisal of personality or for the study of thought processes.

Although it has been generally realized that the choice of stimulus words will have an important bearing upon the results obtained, relatively little has been accomplished in the way of developing an objective criterion for the selection of a suitable list of words for general use, even though the advantages of working with a standard list for which a dependable table of response frequencies has been prepared are well recognized. Undoubtedly, one of the reasons for this seeming neglect is the fact that few people are wholly aware of the importance of a thoroughgoing knowledge of the range of responses to each stimulus word used if errors of interpretation are to be avoided. The distribution of responses for groups varying with respect to such factors as sex, age, education, and social background is likely to differ considerably, even for the most commonly used English words. For this reason, a single frequency table in which the responses of all available subjects are thrown into one common distribution can serve only a limited purpose, but for the same reason a comparison of the results obtained for contrasted groups, selected on the basis of some known differentiating factor, may yield exceedingly valuable information.

If word association tests are to be used for the study of individual or group differences, two criteria for the selection of stimulus words are of prime importance: (1) the words must be simple, falling well within the everyday vocabulary of all those with whom the test is to be used, and (2) each word should tend to elicit responses of strongly contrasted meaning that lend themselves readily to classification. The first criterion can be met reasonably well by comparison with such a standard source as the Thorndike Word List. Some years ago it occurred to me that an excellent approach to the second might be made by confining the choice of stimulus words to homographs—words which, in their written form, have two or more different origins and consequently two or more well-contrasted meanings. Under pressure of instructions to work as rapidly as possible, the subject is rarely conscious of more than one possible meaning for such a word. His choice is made automatically and for the most part without awareness that choice is involved.

Take, for example, the word "ring," which may refer either to an article of jewelry or to the sound of a bell, or, somewhat less commonly, to other circular objects or to abstract shape. In addition, for a small group of persons, chiefly males, the most prominent association has to do with sports such as boxing, while with some women, for the most part those who have a good deal to do with the ailments of children, the word promptly elicits the response "-worm." Literary associations such as "and the book" or "Niebelungen" occur occasionally, and other rare but entirely comprehensible associations will be found in practically any sample of moderately large size.

These are the main categories under which at least 95 per cent of all the responses of an ordinary group of English-speaking subjects can be classified. Subdivisions within these broad groupings can also be distinguished with but slight falling off in reliability of placement. Within the jewelry category, for example, one may note a distinct difference in feeling tone between such comparatively impersonal and matter-of-fact associations as "ring—hand" or "ring—finger" and the more intimate note sounded by "ring—wedding," "ring—engagement," or the still more personal "ring—from Jim." Even more remote than the references to hands or fingers are the not infrequent associations with materials or settings such as "ring—gold" or "ring—sapphire." In like manner associations clearly referring to sound admit a number of subdivisions. There are references to sound in the abstract, as "ring—noise" or "ring—clang." There is the relatively nonspecific "ring—bell." There is the social or business reference, "ring—telephone," or the more exclusively social "ring—door" or "doorbell." There is the definitely active "ring—up" and the more passive "ring—rang."

That the unconscious reading of one or another of these contrasted meanings into a word to which any of them may apply is not merely a circumstantial

matter changing from moment to moment is attested by a number of facts. First, the responses of the same person to each of a series of over 200 words of the homographic type are internally consistent. The consistency is not always apparent to one unacquainted with the underlying pattern of personality and its expression. I doubt that many persons would predict that persons whose response to the word "ring" is "finger" or "hand" are about three times as likely to associate the word "wind" with "rain" as are those whose response to the word "ring" is "bell." I have no very plausible explanation for this, but it has proved to be true in a number of different populations. Neither am I able to say why "ring—finger" and "wind—rain" should be tied in with "kind—good," while "ring—bell" and "wind—storm" are more likely to go with "kind—gentle," but this is also the case. Not only is there this internal consistency but tests repeated after an interval of a year or more show a high degree of correspondence with respect to the classes of response given on the two occasions. Finally, there is the fact that groups selected on the basis of some external criterion exhibit recognizable patterns of response that distinguish them from other groups of different composition.

Observation of this fact led to the hypothesis that free associations to stimulus words of the homographic type may provide unusually rich material for the study of personality differences. The work of Terman and Miles (*10*) on mental masculinity and femininity suggested that a study of sex differences in response to a list of such words would provide a suitable test of the extent to which these unconscious choices among concealed alternatives may be regarded as signs of an underlying pattern of personality structure.

The results of the first study have already been reported (*2*). In this investigation responses to a list of 240 stimulus words of the homographic type from a sample of 800 high school and college students equally divided as to sex were tabulated separately and a weighted scoring key was derived on the basis of the reliability of the sex differences as indicated by the magnitude of Chi-square. By means of this key the responses of an individual subject could be scored for mental masculinity or femininity by a procedure similar to that employed by Terman and Miles.

When judged by the usual criteria, the results obtained by this method appeared at first to be reasonably satisfactory. The sexes were clearly separated with only a small percentage of overlapping. Internal consistency as indicated by the correlation between odd- and even-numbered items was approximately +.80 for single-sex groups and well over +.90 for groups including both males and females. Nevertheless, as data were accumulated it became increasingly evident that something was radically wrong. The nature of the difficulty could be summed up in a few words. The "feminine" man is not the same kind of human being as the feminine woman, nor is "masculinity" in women truly comparable to masculinity in males. There are enough surface resemblances to suggest a name for the trait, but the basic qualities are different. Since this was so evidently true, the use of a common key for scoring the responses of both sexes seemed a questionable practice. A second key or, more precisely, two keys were therefore constructed by the same method used in deriving the first except that the groups compared were the top and bottom quartiles of single-sex distributions of cases which had been scored according to the first key. The two keys thus derived necessarily bear some resemblance to each other and to the first key, but there are also many differences between them which follow a sufficiently consistent pattern to be of some psychological interest.

When reference is made to an activity that is decidedly more common in one sex than in the other, the key for the sex of which the activity is characteristic usually shows a sharp differentiation in the weights given to responses taking the form of active verbs as compared to those for participles ending in *ing* or to the past tenses ending in *ed*. For the sex less often participating, however, no such differentiation in score value usually appears. For example, the stimulus word "rod" elicits from both sexes a good many references to fishing, but in the case of males, those whose total scores fall in the top quartile for masculinity are so much more likely than are those with scores in the lowest quartile to respond with the active verb, "fish," that on the basis of the Chi-square criterion "fish" receives a weight of 4M on this key while "fishing" gets a weight of only 1M. In the case of women and girls, however, no such differentiation between the active verb and the participle appears. Both forms are given somewhat more often by the less feminine than by the more feminine end of the distribution and so receive a weight of 1M, but the sharp cleavage between the active and passive references that is marked for the males is not apparent among females.

When we come to words that refer to a typically feminine activity, the picture is reversed. An example is seen in the word "yarn," which elicits from both sexes a good many references to knitting. Here the sharp distinction between the active verb and the participle is seen in the key used for females for whom the response, "knit," gets a weighting of 4F while "knitting" is weighted only 1F. Again there is no differentiation between the two forms of the verb when given by males. Both receive a score of 1F.

Results such as these strongly suggest that the extent to which an activity has become a part of the life pattern of an individual, welded, as it were, into the very structure of his personality, is indicated not only by the fact that he has formed associations about it but even more strongly by the fact that these associations are related to the concrete performance of the act and not merely to acquaintance with it in a more abstract or remote sense. We fish or we talk about fishing; we knit or we see other people knitting. The one is active, forceful, dynamic; the other is passive, remote, desultory.

Further evidence as to the differential significance of the active verb and the participle is obtained from a second study using the same list of stimulus words. The purpose of this investigation was to ascertain whether or not qualities of leadership are revealed in the free associations of typical adults. Through the cooperation of Col. Hobby the association test was given to approximately 1,000 women in the first Officer Candidate School of the Women's Army Corps, at that time known as the Women's Army Auxiliary Corps. Because of the unusually great care taken in the selection of candidates for this school to secure only women with high leadership ability (*3*), the group chosen seemed likely to provide exceptionally good material for studying this question, at least as far as leadership in women is concerned.

For comparison with the results obtained for the officer candidates, a second sample of approximately 1,000 WAAC privates was also tested. A scoring key for leadership in women was then prepared by the same method as had been used in deriving the M–F keys, using as the basis for deriving the weights the difference between the responses of 500 officer candidates and an equal number of privates matched with the officer sample for age, residence, marital status, and occupation before enlistment.

The proportion of active verbs among the associations given by the officer candidates was many times greater than that given by the privates. This tendency was so marked throughout the entire list of words that it frequently, though not invariably, outweighed the choice of semantic category in determining the leadership score. Tabulation of the weights given in the key indicates that the chances are almost 10 to 1 that a response which is an active verb will be weighted in favor of leadership while, conversely, a response in terms of a participle ending in *ing* has not more than 1 chance in 20 of receiving a positive weight. All this leads to the conclusion that the leader is the doer—the one whose primary associations are with actions and not merely with activity in the abstract.

It was noted earlier that, although the two scoring keys for appraising mental masculinity or femininity in males and in females, respectively, are correlated with each other, they differ markedly in a number of respects. In some cases this difference is so great that the score for the same response may even take opposite directions in the two keys. Of particular interest are those cases in which the score in both keys goes contrary to the sex of the respondent, the same response receiving a feminine weight when given by males and a masculine weight when given by females. An example is "scab—blood," for which women receive a score of 1M and men a score of 1F. Other instances of this class are "kind—mean"; "case—sickness"; "lock—shut"; "can—cannot"; "stick—stuck"; "skip—school" and "skip—miss"; "coast—clear"; "hard—difficult"; "kick—in pants" and "kick—person"; "block—blockhead" and "block—obstruct"; "firm—harsh"; "blow—hard"; "drop—cough." With very few exceptions the list just given appears to suggest some degree of maladjustment.

In marked contrast are the associations for which the weights in both keys correspond to the sex of the subjects. Examples are: "gas—car"; "note—book"; "party—dance"; "case—cokes"; "date—sweetheart" and "date—party"; "dear—mother"; "stage—play"; "strike—baseball"; "coach—athletics"; "dressing—dance"; "cheer—leader." Although many of these responses are rather innocuous, not one appears to suggest the morbidity, resentment, or discontent typical of those in the first list. In general they suggest a healthy preoccupation with the normal activities of young people and a cheerful, contented outlook upon life.

The contrast between the two lists warrants the inference that marked nonconformity with the habits of thinking characteristic of the group to which one belongs is likely to be associated with personal unhappiness or social difficulties. The difference also provides further evidence that the terms "masculine" and "feminine," when applied to the sex which they do not normally characterize, are at most only partially appropriate. The interests and attitudes displayed by the two sexes are too deeply rooted in nurture, if not in nature, to be displaced without risk of damage to the entire personality structure.

Certain trends, however, appear to be characteristic of "masculinity" or "femininity" regardless of the sex of the person who displays them. It will surprise no one to learn that references to mechanics or mechanical processes usually, though not invariably, receive a masculine weighting in both keys while references to domestic activities, clothing, or cosmetics are likely to be weighted in the opposite direction. References to food and particularly to sweets are also much more frequent at the more feminine ends of both the

male and the female distributions. Males more often respond in terms of opposites and adverbial completions than do females while the latter are more inclined to use synonyms and descriptive adjectives. A similar trend appears at the corresponding extremes of both sex distributions.

Interesting, also, is the correlation between the total M–F score and the number of overt references to people. We have classified under this head only such nouns as clearly refer to human beings, such as "man," "woman," "blacksmith," "person," and so on, together with their plurals. In 1922 Moore (*7*) called attention to the greater frequency of references to persons and personalities by women as compared to men found in the snatches of talk overheard by an observer stationed at a busy street corner in New York City. A similar procedure was followed by Carlson and his associates (*1*) in the lobby of the concert hall during intermissions of the Minneapolis Symphony Concerts with similar results. McConnon (*5*) carried the procedure to the five-year-old level, using the free-play period in a number of Minneapolis kindergartens. Although the sex difference in the number of remarks having to do with people was less marked than those established for adults, it was in the same direction and was large enough to be significant at the 1 per cent level of confidence.

In all these studies the comparison was made between individuals who actually differed as to sex. It is the more interesting, therefore, to find that the free associations of persons of the same sex but differing with respect to the M–F score obtained on their total list of responses show a similar trend. For example, when a count was made of the number of nouns referring to human beings among the responses of a group of 100 adult women and the results tabulated according to the total M–F score, the means for the successive quartiles from the most to the least "feminine" were as follows: top quartile, 17.5; remaining three, 10.5, 7.5, and 6.0. Even more striking are the figures for extreme counts. In the most "feminine" quartile, 37.5 per cent had counts as high as 20 while no one in either of the two lowest quartiles reached this figure. Of those in the most "feminine" quartile, only 25 per cent gave fewer than 10 responses of this class while in the lowest quartile 96 per cent of the scores were below 10.

Tabulations for a group of male subjects showed a similar trend, although the difference was less marked. The comparative lack of discrimination in males arises from the small number of very high scores, which reduces variability at the upper extreme, while the fact that a few words in the list elicit references to human beings in a large percentage of cases of both sexes makes it rather unlikely that any score will fall below 3 or 4. However, in spite of the marked reduction in range, it was found that means for successive quartiles arranged in order from highest to lowest masculinity were as follows: 8.6, 10.1, 10.9, and 12.4. In this group only 2 of the 100 subjects had scores as high as 20.

The relation of masculinity or femininity to factors in the life of the subjects is a fascinating but complicated problem. A number of intrafamily relationships have been noted, and other studies in this field are in progress. A recent study by Smith (*8*), in which the association test was given to a group of university girls and their parents, yielded correlations of +.26 between scores of mothers and daughters and −.33 between those of fathers and daughters—that is, there was a small but reliable tendency for the most feminine girls to have the most feminine mothers and the most masculine fathers. There was also a slight tendency, at least within this group, for the most masculine husbands to have the most feminine wives. Whether this relationship is a matter of marital selection or has arisen from the give and take of marital life cannot be determined from the data at hand, since most of the couples had been married to each other for around a quarter of a century at the time the test was taken. When a midparent score was obtained by averaging the scores of husband and wife without regard to sign, the correlation with daughters' scores was +.40.

Using the M–F key derived from the Strong Vocational Interest Blank, Terman and Buttenwieser (*9*) made a special study of the masculinity–femininity scores of happily and unhappily married couples and of couples who had been divorced. An outstanding finding was the very high masculinity score of the divorced women. I have compared the scores on the association test of divorced, married, and single women, matched for education, age, and occupation. One hundred cases were included within each group. The range of ages was 21 to 44 years. For purposes of comparison the cases were divided into two age groups of 50 sets each, those of 32 years or older constituting the older group and those under 32 comprising the younger group.

Among the older group the number of divorced women making definitely masculine scores was reliably greater than that for either the single or the married women, the difference being significant at the 2 per cent level of confidence in both cases. Among the younger women the difference was in the same direction but not large enough to be reliably established. Possibly the fact that more subsequent changes in marital status are to be expected among this group accounts for the smaller degree of differentiation.

It is interesting to note that the group of divorced

women also includes a greater number of cases with very high feminine scores than either of the other groups, in addition to the excess of masculine scores. The difference is small but in the same direction at both age levels. Possibly the woman whose interests are too exclusively of the feminine pattern, who has not enough admixture of masculine interests to permit real companionship, may be as poor a risk in marriage as the one who competes with her husband for the masculine role.

That stability in marriage may be an indication of stability in a broader sense is suggested by the definitely smaller number of rare responses given by the married women. Practically all investigators have agreed that the occurrence of an exceptionally large number of unusual responses is frequently associated with some form of mental or emotional disturbance. We have classed as "rare responses" those given by fewer than .5 per cent of the 800 cases used for standardization. Tabulation of the frequency of such responses for the three marital groups shows a marked preponderance of very high scores among the single women and the divorced, with relatively few among the married group. Whatever may be the factors lying back of these high scores, they at least seem to be much less frequent among married women than among those who have either failed to marry or whose marriages have been unsuccessful.

The past few years have witnessed a striking shift in the techniques employed for the appraisal of personality differences. Earlier studies in this area were dominated almost completely by the theories and methods that have proved so useful in the measurement of ability. For the most part these are sampling devices. Gradually we have come to realize that there is a fundamental difference between ability and attitude which is of basic consequence for the problem of measurement. In the case of ability, the universe to be appraised is, as a rule, overt and readily accessible. Few people insist upon hiding their talents and skills in a napkin. The method of sampling is therefore well suited to the purpose at hand, the major problems arising from lack of agreement among workers as to the limits of the universe and uncertainty as to the best means of obtaining an adequate and representative sample of it. But when we come to deal with what is often called the "private world" of the individual, comprising, as it does, the feelings, urges, beliefs, attitudes, and desires of which he may be only dimly aware and which he is often reluctant to admit even to himself, much less to others, the problems of measurement are of a very different nature. Here the universe which we wish to assay is no longer overt and accessible but covert and jealously guarded. Thus, the methods of direct sampling become hazardous since an unbiased sample is so difficult to secure.

Recognition of this difficulty has led modern students of personality to a search for signs rather than samples. The devices employed in the hope of eliciting signs by which the inner life of the individual may be unconsciously revealed are often grouped together under the heading of "projective methods." The Rorschach ink blots and the Thematic Apperception Test are among the better known examples of this class. An outstanding advantage of these methods is the fact that, because the task which is set the subject is relatively unstructured, he is free to respond in whatever way seems to him to be most nearly right and meaningful. Since no set scale of answers is provided, the only suggestions given are those arising from his own individual interpretation of the situation in the light of his experiences and attitudes. While this freedom of expression makes for many difficulties in scoring, the time and effort expended should be well worth while in view of the greater potential value of the material obtained.

References

1. CARLSON, J. S., COOKE, S., and STROMBERG, E. *J. appl. Psychol.*, 1936, **20**, 727–735.
2. GOODENOUGH, FLORENCE L. The use of free association in the objective measurement of personality. In *Studies in personality contributed in honor of Lewis M. Terman.* New York: McGraw-Hill, 1942. Chap. 5.
3. GOODENOUGH, FLORENCE L. *Psychol. Bull.*, 1942, **39**, 634–637.
4. KENT, G., and ROSANOFF, A. J. *Amer. J. Insanity*, 1910, **67**, 48–96.
5. MCCONNON, ALICE. Sex differences in topics of conversation of kindergarten children. Unpublished Master's thesis, Univ. Minnesota, 1938.
6. MALLER, J. B. *A controlled association test.* New York: Bureau of Publications, Teachers College, Columbia Univ., 1934.
7. MOORE, H. T. *J. abnorm. Psychol.*, 1922, **1**, 210–214.
8. SMITH, JANE HARRIS. *J. soc. Psychol.*, 1945, **22**, 79–85.
9. TERMAN, L. M., and BUTTENWIESER, PAUL. *J. soc. Psychol.*, 1935, **6**, 143–171; 267–289.
10. TERMAN, L. M., and MILES, CATHERINE COX. *Sex and personality.* New York: McGraw-Hill, 1936.
11. WYMAN, JENNIE B. Tests of intellectual, social, and activity interests. In L. M. Terman. *Genetic studies of genius.* Vol. I: *Mental and physical traits of a thousand gifted children.* Stanford Univ.: Stanford Univ. Press, 1925. Chap. 16.