In personality theory a ubiquitous and fundamental distinction may be drawn between the interpretation of behavior in terms of (a) the content of "needs" and of cognitive structures generally and in terms of (b) characteristic styles of response and action. The separation of these two components of personality organization has taken a variety of forms in the hands of different theorists, as in the Allport-Vernon (2) Studies in Expressive Movements, in Murphy's (47) scholarly discussion of continuity in personality structure, in Klein's (40) distinction between needs and control processes, and in Vernon's (54) distinction between adaptive and expressive behavior. One may legitimately ask not only what a person says or does (the particular content of his statements and actions) but how he acts (his characteristic mode or style of expression).

What is conceptually a relatively sharp distinction is typically blurred and confounded in a particular concrete act; the what and how are fused in a given goal-directed response. An obsequious person indicates his deference not only by the act of yielding, but by the tone of his voice in performing the yielding act. Because content and style are intermixed in a given behavior sequence, and because there is often a theoretical predilection for content components, style is often overlooked in personality assessment. Also, the measurement of content appears to be more direct and unambiguous than the assessment of stylistic dimensions of personality. It is possible, for example, to ask a person what his attitude is on a given topic, or to draw inferences about his need patterns from his reported likes and dislikes (51). The obviousness of such devices, while helpful from the viewpoint of labeling what one hopes one is measuring, also permits respondents to distort their scores if they so desire (32), something which is less likely to occur in the assessment of style.

In considering the general distinction between content and style, those methods of personality and attitude assessment which are based upon printed questionnaires of one form or another will be emphasized. While the complementary constructs of content and style have special relevance to questionnaire items, where the response-evoking properties of the particular item form may contribute markedly to response variance above and beyond the contribution of content, the distinction might also be applied usefully to other areas of personality assessment. For example, three possible applications are to perceptual and cognitive style as in the work of Thurstone (52), Witkin (58), Klein (39, 40), Gardner (21), and others (34); to achievement and apti-
tude testing (28, 32, 60); and to the perception of personality (2, 38, 54, 59).

The present discussion attempts to do two things: first, to present some evidence showing the important and subtle influences upon responses of stylistic components of item form; and, second, to illustrate how reliable measures of potentially useful stylistic dimensions may be generated from characteristic responses to the form of personality and attitude items as distinct from measures of content.

**PERSONALITY STYLE AND RESPONSE SET**

Traditionally, responses to a particular item or set of items are assumed to provide information about the respondent in terms of the item content. If, for example, a person agrees with the statement, “Under no conditions is war justified,” or answers “true” to the item, “I have more trouble concentrating than others seem to have,” it is commonly assumed that these responses, if consistent, will indicate respectively something about the person’s attitude toward war or his mental state. Under these conditions response determinants such as the subjects’ generalized tendency to agree are legitimately considered as sources of cumulative error, Cronbach’s (13, 14) familiar “response sets.” While Cronbach’s emphasis was that response sets often lead to errors of interpretation in the logical validity of tests, he also indicated that these response tendencies might not always be temporary and trivial, but may have a stable and valid component which reflects a consistent individual style or personality trait. While recognizing Cronbach’s contribution in describing the phenomenon, it is preferable for the present purposes to change the label from “response set” to components of style. This change in terms emphasizes the fact that for certain purposes in personality assessment opportunities for the expression of personal modes for responding should be enhanced and capitalized upon, rather than considered as sources of error to be avoided or minimized. This change also avoids the ambiguity inherent in the concept of “set” (22).

**CHARACTERISTIC STYLES IN PERSONALITY AND ATTITUDE QUESTIONNAIRES**

Among the more prominent response styles usually evoked by questionnaire items are response acquiescence, overgeneralization, a tendency to respond in a socially desirable way, and the complementary tendencies to respond negativistically, critically, and in a socially undesirable or idiosyncratic manner. Some pertinent illustrations will be drawn of how each of these, operating singly and in combination, may influence the interpretation of responses to psychological tests. Alternative procedures for evaluating these stylistic variables will then be discussed.

**Response Acquiescence and Authoritarianism**

It has been long recognized that a subject who agrees with a personality or attitude item stated in a positive form may not necessarily disagree with its logical opposite, but may instead show a fairly general tendency toward agreement or disagreement. Studies by Rundquist and Sletto (49), by Lorge (42), and reviews by Cronbach (13, 14), Berg (8), and Messick and Jackson (45), indicate that response acquiescence is widespread and pervasive over a wide
variety of item content and most pronounced when content is highly ambiguous or imaginary. Berg (8, 9) has suggested that acquiescence is a modal response in our culture when the issue before the respondents is unimportant or nonexistent.

The operation of such stylistic tendencies should be taken into account in the course of personality measurement. If a particular content area is to be assessed, it is at least necessary to introduce into the scaling procedure appropriate experimental controls for acquiescence, or else reconcile oneself to interpretive equivocality due to the confounding of content and style in a single measure. Other response determinants besides acquiescence, however, must be controlled before characteristics may be unequivocally attributed to respondents on the basis of item content. Messick and Jackson (45) have discussed alternative methods for reducing this ambiguity of interpretation in the measurement of authoritarian attitudes.

Even though much of the recent research with the California F scale (1) has been of a methodological and critical nature, it nevertheless yields some important information on the relationship between content and style. A number of investigators (5, 46, 36, 37, 41, 45) have independently correlated scores based on the California F scale, in which all of the items are so worded that agreement is always scored in the authoritarian direction, with scores based on logically reversed F-scale items. These correlations were not found to be high and negative, as would be expected from consistent responses to item content. With one reversed F scale (36), significant positive correlations in the acquiescence rather than the content direction were obtained. Furthermore, there is evidence (37) that previously obtained relationships between personality variables and the F scale, formerly thought to be interpretable in terms of correlates of authoritarian ideology or content, may need reinterpretation in terms of consistencies in style. The most recent study requiring such reinterpretation is one by Gilbert and Levinson (23), in which a scale purportedly measuring "custodial mental illness ideology" was constructed, with 17 of 20 items requiring agreement to be scored as "custodial ideology." A high correlation between the "custodial ideology" scale and the F scale was used to support the conclusion that "preference for a custodialistic orientation is part of a broader pattern of personal authoritarianism." But Howard and Sommer in a replication found that "custodialism" correlated significantly with agreements to both the original and the Jackson-Messick (36) reversed F scales, which would seem to indicate that style rather than content is of primary importance in this instance. Christie, Havel, and Seidenberg (12) have shown that it is possible in some samples to obtain a correlation

Gage, Leavitt, and Stone (20) have argued that confounding content and style in the F scale, far from being a source of error, is fortunate, because acquiescence contributes to the empirical validity of the F scale as assessed by independent ratings of authoritarian behavior. If the aim is merely to predict authoritarianism as a criterion, like predicting the success of salesmen, this argument might be legitimate as long as the criterion did not change. But if one hopes to understand the various components of a dynamic construct like authoritarianism, conglomerate indices containing both content and style will not suffice and will confuse the issues (45).
between reversed and original F-scale items significant in the content direction. Jackson, Messick, and Solley (37) had previously reported a correlation of +.35 between agreements to original and to reversed F-scale items. What accounts for this apparently considerable discrepancy? One set of investigators predicted and obtained a correlation significant in the acquiescence direction, while another, with a different reversed F scale, predicted and obtained a correlation in the content direction. The answer to this question requires a consideration of more than differences in the content of the two reversed F scales; the form of the items must be examined. Jackson and Messick (36) indicated that the original, extremely worded, cliché-ridden style of the F scale was retained in their reversals, while Christie, Havel, and Seidenberg (12) explicitly avoided the sweeping generalizations found in the originals and substituted much more cautious statements. It is likely that this difference in item form accounts for the different results of the two sets of investigators. It appears that the tendency to endorse statements containing phrases such as “every person,” “no person,” “all,” “most important,” “complete certainty,” “never,” “must,” etc., is a general one, which may act independently of the content. This response style to overgeneralize may contribute to relationships between the F scale and cognitive variables like rigidity (37) and perceptual intolerance for ambiguity (18). It probably also partially accounts for the frequent observation that verbally elicited ethnic attitudes tend to be highly intercorrelated (10), even, for example, in Hartley’s (30) study where the “groups” were nonexistent and no previous attitude or “cognitive structure” could be assumed to exist. An appraisal of variance associated with aspects of authoritarian content on one hand, and stylistic components like response acquiescence and overgeneralization on the other, would seem to require at least four sets of items: an extremely worded original and reversed F scale and a probabilistic original and reversed F scale. It is suspected that subjects endorsing probabilistic F-scale items would not show as much of the “authoritarian’s” intolerance for ambiguity as might be expected, although some relationship between authoritarian ideology and response style might still be obtained.

Response Acquiescence in Personality Inventories

The distinct roles of content and style should also be noted in responses to personality inventories, especially those “true-false” devices like the MMPI developed by the empirical selection of discriminating items. While few, if any, investigators have ever explicitly assumed that the total number of empirically derived scales was the most parsimonious way of summarizing the common variance of an inventory, the use of a large number of separate scales as, for example, in the 9 clinical scales of the MMPI or the 18 scales of Gough’s California Psychological Inventory, is justified by the extent to which each makes some independent contribution to the assessment problem not made by the other scales.4 If there is a great deal

4 The MMPI was advanced initially as an aid in the prediction of psychiatric diagnoses. In practice it is rarely so used in any literal sense, which is fortunate, as the research evidence (e.g., 7, 48) indicates that predictions of specific diagnoses generally cannot be made with certainty. Rather, the original purpose of the MMPI, prediction, has come to be modified so that now scores, singly or in combination, are used to draw inferences about
of common variance among the various scales, this redundancy limits their efficiency.

There is considerable evidence that a very few factors account for the major proportion of the variance on personality inventories of the "true-false" variety. Wheeler, Little, and Lehner (57), for example, reported a factor analysis of MMPI scales in which only two major factors and one minor factor were identified. In the light of accumulating evidence it seems likely that the major common factors in personality inventories of the true-false or agree-disagree type, such as the MMPI and the California Psychological Inventory, are interpretable primarily in terms of style rather than specific item content.

One line of departure from which it is possible to evaluate the role of acquiescence in personality inventories is to consider the percentage of items keyed "true" in each scale as an index of the extent to which that scale elicits response acquiescence. Jackson (33) did this with the California Psychological Inventory, computing rank order correlations between the percentage "true" in each scale and the scale's correlation with outside personality measures shown previously to reflect acquiescence. A number of high and significant correlations with such unidirectional scales as the California F scale and the MMPI K scale suggests strongly that acquiescence is a major source of variance in the CPI.

Messick and Jackson have obtained evidence of a similar nature for characteristics of respondents (56). Somewhat different notions of validity (15) and a different mathematical model (27, 53) are necessary in the latter case.

Messick, S. J., & Jackson, D. N. "Response Style and Factorial Interpretation of the MMPI." In preparation.
sweeping generalizations, it may be that acquiescence on the MMPI is elicited differentially by certain content categories, or in relation with another stylistic component.

The specific source of the variables which appear to moderate the operation of response acquiescence in the MMPI is obviously a complicated research problem which awaits more evidence for a definitive answer. One very promising lead, however, is encountered in another important stylistic determinant of test-taking behavior, the general tendency to endorse socially desirable or socially undesirable statements about oneself. This stylistic response tendency on the part of individuals should be distinguished from the judged characteristics of desirable and undesirable item content. There is considerable evidence that this tendency is general and is related to a tendency to respond in an idiosyncratic or atypical manner. Edwards (16) has reported a correlation of .87 between judged social desirability scale values and the proportion of respondents independently endorsing them. Hanley (29) obtained correlations of .82 and .89 respectively between probability of endorsement and social desirability ratings for samples of items from the MMPI D and Sc scales. Fordyce (17) correlated with the MMPI clinical scales a set of MMPI items judged to be socially desirable. His obtained correlations were high, ranging from −.38 to −.91. Although these coefficients indicate the importance of social desirability in scales like the MMPI, they also reflect the influence of response acquiescence, since the social desirability scale contained a disproportionate number of items keyed false. Jackson (33) showed that a combination of ranked indices of response acquiescence and social desirability on scales of the California Psychological Inventory was related to the rank of each scale's correlation with the MMPI K scale to the extent of \( r = .86 \). This value was higher than the correlation of either response style operating singly, suggesting the possibility of summative effects of response acquiescence and social desirability.

Berg (8, 9), granting that there are modal response patterns, suggested that individual differences, particularly deviations, may be revealing of personality style. Berg hypothesized that deviant behavior tends to be general and not specific to any particular content area. Barnes (3, 4) appraising the Berg deviation hypothesis in the MMPI, shed important light on the relation between an acquiescent style and idiosyncratic responses. Barnes demonstrated a close correspondence between Wheeler, Little, and Lehner's (57) first or "psychotic" factor and total number of items answered deviantly true, and between their second or "neurotic" factor and total number of items answered deviantly false. Although response acquiescence and the tendency to respond in a socially undesirable or deviant manner are confounded in Barnes' analysis, these results strongly support the notion that items judged low in social desirability evoke different tendencies toward acquiescence, as compared with items judged high in social desirability. This interpretation appears consistent with Welsh's (55) data, where the first pure factor scale, composed of 38 "true" items out of 39, contains many socially undesirable statements, while the second pure factor scale, where all the items are keyed false, seems to consist predominantly of neutral or somewhat socially desirable statements. Here again, a consistent response style to acquiescence seems to be elicited differentially by
a variety of self-deprecatory statements on the one hand, while, alternatively, neutral or mildly socially desirable statements evoke consistent differential tendencies to disagree or to be negativistic.

Whether there are consistencies attributable to content after allowing for style in these first two factors or, indeed, in any obtained scores on the present form of the MMPI is an important research question, as is the relation between various content and stylistic factors and psychopathology. If Berg (8) is correct, if one might just as well use abstract drawings (3) as items to discriminate empirically psychiatric patients from normals, then it may be that content is less important and style more important than previously supposed. If this is the case, then past attempts to draw conclusions about respondents on the basis of their answers to uncontrolled item content are suspect. If, on the other hand, consistencies in content can be demonstrated above and beyond components of style, it is extremely important that measures of these content variables make adequate use of proper experimental controls to avoid as far as possible confounding with style. Use of recent advances in scaling theory (27, 53) might be helpful.

Measuring Personality Styles

In approaching the problem of the assessment of style, a curious dilemma presents itself. On the one hand, it is easy to show that most personality tests are loaded with stylistic components, but on the other hand, good measuring devices for these dimensions do not exist, largely because few research workers have attempted explicitly to devise such scales. Typically, a single measure, like the California F scale, the MMPI K scale, or Bass’s (6) collection of aphorisms, has been offered as an index of a response style, acquiescence, for example. Little thought is given to the fact that these measures may not only contain several dimensions of content, but of style as well, thus limiting their usefulness as indices of any particular style. Thus, Fordyce (17) has suggested that the MMPI K scale reflects tendencies to respond in a socially desirable manner, while Fricke (19) has argued that the K scale reflects acquiescence. Evidence from each of the two authors is convincing, and, indeed, Jackson’s study (33) supports the notion that the K scale contains both acquiescence and social desirability variance. It may reflect other things as well, but this confounding is not conducive to its use as a measure of one particular style. The same criticism might be leveled at the California F scale, at Edwards’ (16) social desirability scale, and at Bass’s (6) social acquiescence scale, all of which seem to confound response acquiescence with social desirability.

One way to construct measures of such styles as acquiescence or overgeneralization would involve selection of items extremely heterogeneous in content. Experimentally independent measures of each style would, of course, be desirable. Since a response style to answer in a socially desirable or undesirable direction seems to be omnipresent, it is hard to avoid in measures of other styles. Rather than attempting to develop items all at one level of social desirability, it might be better to vary social desirability systematically and to observe its relationships and interactions with other variables. Helmstadter (31) has described procedures for obtaining separate scores for different components of a test, some of which would be especially relevant to a situation in which one
had already obtained social desirability scale values. Although social desirability has been assumed to be one-dimensional, it is easy to conceive of distinct, but perhaps correlated, dimensions consisting of items reflecting irresponsibility, psychiatric bizarreness, or hostility. The selection of sets of items for different dimensions of judged social desirability would be facilitated by the application of recent advances in multidimensional scaling (44). Such refinements as separating out the components of social desirability would do much to clarify response determinants and might put personality evaluation upon a more rigorous basis than has previously been thought possible.

Although the emphasis in this paper has been on some of the more conspicuous stylistic determinants encountered in common personality tests, there are many other possible measures of style that might be derived from personality theory. For example, a tendency to express a liking for diverse things, although it might be response acquiescence in a new disguise, might also represent greater cognitive differentiation or capacity to invest energy freely in objects in one's environment. Such general expressions of "like" and "dislike" have been found to be reliable. On one set of 300 items dealing with diverse activities (51), the corrected split-half reliability of the tendency to respond "like" was .86. With a paucity of evidence on these issues, the alternative to such conjecture is carefully planned research, for which there is an obvious need. There are many other research opportunities for the measurement of style, such as asking respondents to select from among two or more personality, attitude, or achievement items, equal in valence or correctness, but perhaps couched in different phrasings—perhaps one elaborate and pedantic, one simple, and one containing slang. Preferred modes or styles of expression might also be readily evaluated by techniques disguised as achievement tests (32). In this context, it would be interesting to evaluate personality correlates of such attributes as tolerance for logical contradictions within a passage, of a tendency to gamble on achievement tests (28, 54), and a variety of other consistent modes of response. Similarly, further research is needed to evaluate Jackson's (34, 35) hypothesis that respondents who acquiesce consistently manifest a lower level of cognitive energy in other situations.

**Summary**

It has been suggested that stylistic determinants, such as acquiescence, overgeneralization, and a tendency to respond in a socially undesirable, overgeneralized manner, as distinct from specific content, account for a large proportion of response variance on some personality scales, particularly the California F scale, the MMPI, and the California Psychological Inventory. In developing and evaluating measures of style it is important to select not only those measures which have appeared by accident on already established tests, but to design assessment techniques explicitly to evoke theoretically important styles of response. Research involving response style may contribute to a more systematic measurement in personality and may pay off handsomely in helping to further the common ground between personality theory and personality assessment.
REFERENCES


16. EDWARDS, A. L. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. J. appl. Psychol., 1953, 37, 90-93.


33. JACKSON, D. N. Response acquiescence in the California Psychological Inven-
Received November 14, 1957.