How relevant is an individual difference theory of personality? ¹

Richard A. Shweder, ² Committee on Human Development, University of Chicago

A type of explication of the concept "personality" having wide currency in the personality literature refers to "more or less stable internal factors that make one person's behavior consistent from one time to another, and different from the behavior other people would manifest in comparable situations" (Child, 1968, p 83). In this study, the degree of relevance of the concept "personality" construed in this "individual difference" sense is questioned. Some of the most ubiquitous personality assessment procedures, specifically interpersonal checklists, personality inventories, and questionnaire interviews are shown to be subject to a form of systematic bias which creates the "illusion" that the diversity of individual behavioral differences can be reduced to a limited set of person-distinguishing underlying forces (traits, dispositions, scales, factors or dimensions).

Respondents on interpersonal checklists, personality inventories, and questionnaire interviews unwittingly substitute a theory of conceptual likenesses for a description of behavioral co-occurrences. Propositions about theory are confused with propositions about the world (D'Andrade, 1965, p 215) and con-

¹ An earlier version of this manuscript entitled "Illusory Correlation and Personality Theory" was prepared for the Mathematical Social Science Board Conference "Field Methods for the Collection of Verbal Data," Quail Roost, North Carolina, May 2-4, 1974, which in turn was an expansion of one section of a doctoral thesis Semantic Structures and Personality Assessment submitted to the Department of Social Relations, Harvard University. I wish to thank the National Institute of Mental Health (fellowship #1 FO1 MH 43845-01 CUAN) and the Spencer Foundation for financial support, and a number of scholars for their encouragement and / or criticism, in particular, R. Freed Bales, Jerome S. Bruner, Donald W. Fiske, Walter Mischel, John W. M. Whiting, and especially Roy G. D'Andrade. Nothing concerning their degree of agreement with the views set forth in this article is implied by this expression of appreciation.

² Reprint requests should be sent to Richard A. Shweder, University of Chicago, Committee on Human Development, 5730 S. Woodlawn Avenue, Chicago, Illinois 60637.
siderations about similarity confounded with judgments about probability (Chapman, 1967, Chapman and Chapman, 1967, 1969, Tversky and Kahneman, 1974) Items alike in concept are inferred to be behaviorally characteristic of the same person even when, as is typically the case, conceptual relationships among items do not correspond to the actual behavioral relationships among items. As a consequence respondent judgments about "what goes with what" on checklists, inventories, and questionnaire interviews are systematically distorted in the direction of pre-existing conceptual associations concerning "what is like what" which are themselves not inductive generalizations from actual behavioral instances. These conceptually biased judgments create an "illusion" of underlying individual behavioral consistency, which, although not apparent in actual behavior, deceptively validates the "individual difference" conceptualization of "personality."

The adequacy of the "personality" concept in its "individual difference" sense has been assessed almost exclusively in relationship to judgmental data from interpersonal checklists, personality inventories and questionnaire interviews. The personality literature is replete with checklists, inventories and questionnaires from which traits, dispositions, scales, underlying states, factors and dimensions for describing individual differences have been statistically constructed, and success at such scale and factor construction has been interpreted as prima facie support for the applicability of the "personality" concept. Clinical psychologists concerned with psychopathology speak of persons who are "ego-resistant" and point to a pattern of associated scales on the MMPI (e.g., Block, 1965) Social psychologists concerned with interpersonal behavior in groups speak of persons who are "dominant" and point to a pattern of associated items on an interpersonal rating form (e.g., Bales, 1970) Child psychologists concerned with child rearing practices speak of mothers who are "permissive" and point to a pattern of associated scales on a socialization questionnaire (e.g., Sears, Maccoby and Levin, 1957)

In the last decade, however, a potentially devastating methodological critique of exactly this type of judgmental data has emerged. Focusing upon the types of inferences legitimately
drawn from certain kinds of verbal reports about behavior, Mulaik (1964) and D'Andrade (1965) draw a logical distinction between "how people classify" and "how to classify people," between conceptual likeness and behavioral co-occurrence. They demonstrate that the factors underlying interpersonal trait ratings can be independently derived from purely conceptual criteria, e.g., judgments about the similarity of meaning of trait words. Rating factors can be discovered by simply asking raters to judge how conceptually similar trait words are to each other without any reference to actual behavior.

Mulaik and D'Andrade are divided over the implications of their findings for questions having to do with the degree of relevance of the "personality" concept in its "individual difference" sense. Mulaik believes there must be some systematic fit between conceptual factors and factors which might be obtained from immediate records of behavior. D'Andrade disagrees. He claims that not only have raters and psychologists logically confused "propositions about the world with propositions about language" (i.e., person factors with stimulus factors) but also that the factors discovered in ratings are the property of trait terms as linguistic elements without being the property of the persons described by those terms. Nonequivalent conceptual-logical relations have been unwittingly substituted for factual-empirical relations in the process of making a judgment. The factors discovered in ratings, he argues, tell us "how people classify" (and by implication "that people classify") and no more. They should not tell us "how to classify people" (or "whether to classify people").

There are three sensible interpretations compatible with Mulaik's and D'Andrade's finding that rating factors can be derived from conceptual criteria:

1. There is a spurious relationship between conceptual factors (x) and rating factors (y). Our conceptual associations (x) and our judgments of "what goes with what" in behavior (y) are independently sensitive to the same true personality factors (z). i.e., our judgments about behavior are inductive generalizations from experienced instances, and our language (presumably over a considerable period of time) has encoded these same generaliza-
tions about behavioral regularities which we learn through observation. Z causes x and z causes y.

2 There is a casual non-distorting relationship between conceptual factors (x) and rating factors (y). Conceptual factors (x) control judgments (y) but reflect true personality factors (z), i.e., our judgments about behavior rely on pre-existing conceptual associations among trait labels but these are accurate reflections of behavioral regularities. Z causes x which in turn causes y.

3 There is a causal distorting relationship between conceptual factors (x) and rating factors (y). Conceptual factors (x) control judgment (y) but are not related to true personality factors (z), if such factors exist at all. X causes y and is unrelated to z.

The overwhelming consensus of psychological opinion, including Mulaik’s favors one of the first two interpretations, preserving as do these two interpretations the existence of true personality factors, and the relevance of the “personality” concept in its “individual difference” sense for an account of behavior. For example, Bruner, Shapiro and Taggiuri (1958, p. 278) argue that trait terms and the inferences we make from one trait to another form part of a “coding system” about the nature of persons, support for which resides in the very consistency of behavior. Trait words, they believe, “summarize” or “package” these consistencies in behavior and are thus more or less accurate. Passini and Norman (1966, p. 47) argue that “in the course of observing and interacting with other persons” implicit personality theories are developed which encode the actual “relative frequencies of joint occurrences of various personality attributes and behavioral dispositions in other persons.” Roger Brown (1965, p. 612) argues that trait words (e.g., “friendly”) are summary formulas about person-dispositions which generate expectations which in turn have utility in predicting future events and minimizing surprise. “From hearsay and direct observation we have learned that certain kinds of behavior hang together, not invariably but probabilistically. People who like parties also introduce themselves to strangers and stop to chat with friends and smile easily and so on.”

D’Andrade was bold in his original interpretation that the
interpersonal rating factors used to distinguish persons were of “illusory validity.” He was expressing an unpopular opinion on the basis of equivocal data. Recent evidence has nonetheless lent credence to D'Andrade's position. Chapman and Chapman (Chapman, 1967, Chapman and Chapman, 1967, 1969) have shown that the pre-existing verbal association strengths between pairs of clinical diagnoses and psycho-diagnostic test items systematically bias estimates about the frequency with which those pairings co-occur in spite of contradictory observational evidence (also see Tversky and Kahneman, 1974) and D'Andrade himself in a study of interpersonal behavior in small groups (1970) has demonstrated that judgments about interpersonal behavior systematically move in the direction of the judge's pre-existing understanding of “what is like what.” The more conceptually similar two persons' attributes (trait labels) are, the more likely they will be judged characteristic of the same person. This is true even when it can be demonstrated that in the judge's past experience the two attributes were not in this manner associated. The outcome of such biased estimates of the co-occurrence likelihoods of pairs of events is an “illusory correlation” (Chapman, 1967). A number of observers agree that a correlation exists which is not warranted by their own experience.

It is the purpose of the present study to support D'Andrade's position with additional evidence and to extend the thrust of his critique from interpersonal ratings to other personality assessment procedures, including personality inventories and questionnaire interviews. Respondent judgments of “what goes with what” will be shown to replicate pre-existing understandings of “what is like what” despite contradictory observational evidence. Pre-existing understandings of “what is like what” will be shown not to necessarily reflect the co-occurrence relationships among actual behaviors. Finally, evidence generally thought to demonstrate the applicability of the “individual difference” conceptualization of “personality” will be shown to be equivocal by deriving the trait-factors discovered in interpersonal rating forms, questionnaire interviews and personality inventories from purely conceptual criteria without any reference to actual behavior.

If this critique in its extended form is substantially justified then verbal data collecting procedures, such as interpersonal ratings, self-report inventories, and retrospective questionnaires,
are more appropriately viewed as techniques for giving us access to cognition than behavior, in this case, to a construction of the nature of the "person" (undeniably of enormous interest from a cultural or attributional point of view) which cannot be systematically supported by detailed and immediately recorded observational data.

**Experiment I**  **The Conceptual Determination of “Illusory” Rating Factors**

**Method**

Newcomb's 1929 study of extroversion-introversion is in many ways the forerunner of diverse contemporary critiques of theory and method in personality psychology (e.g., Burton, 1970, Campbell & Fiske, 1959, D'Andrade, 1965, 1970, Fiske, 1974, Mischel, 1968, 1973, Shweder, 1972, Yarrow, Campbell & Burton, 1970) The lack of consistency of behavior, the questionable accuracy of retrospective report, the absence of convergent and discriminant validation for personality constructs, and the influence of conceptual presuppositions on judgment are all noted in what must be considered a forgotten classic of psychology Some of Newcomb's data shall be reanalyzed below

Newcomb's data consists of daily records and observer ratings of the same 26 items of behavior for two groups of boys in a summer camp (N = 27 and 24) Each group of boys was observed for 24 days in various camp settings The two groups of boys differed in age Daily records on each boy were kept by a counselor who noted the presence of an item of behavior as soon after its occurrence as possible An odd-even day reliability check on behavior percentages produced a mean reliability coefficient of .78 over the 26 items of behavior Six observers, including the counselor, rated the boys at the end of the camp session on a 5-point scale for each of the 26 items Their ratings were pooled in subsequent analyses The 26 behaviors used in Newcomb's study are

1. Speaks with confidence of his own abilities
2. Takes the initiative in organizing games
3. Submits quietly and with good spirit to criticism or discipline from his counselors
4. Does more than his share of work for inspection
5. Gets up before rising hour
6. Willingly falls in with what others in his group want to do
7. Speaks or laughs aloud when the group is supposed to be quiet
Individual difference theory of personality

10 Eager to be first when the group takes turns or gets in line
11 Engages in group misdemeanor
12 Spends more than an hour of the day alone
13 Takes the initiative in approaching or speaking to a stranger
15 Tells of his own past or of exploits he has accomplished
16 Gets into scraps with other boys
18 Hesitates to take his turn at serving at the table
19 Painstaking in making his bed
20 Reads a half-hour or more during the day
21 Spends swimming periods in the water, actively moving about
22 Works steadily, without stopping, at after-meal work
23 Gives loud and spontaneous expressions of delight or disapproval
24 Gets into trouble of a mischievous or adventurous nature
25 Conversations with counselors confined to asking and answering necessary questions
26 Uses exaggerated gestures, antics, or show-off activities
27 Spends quiet hour without loud talking or running about
28 Continues on a single activity for the whole morning
29 Talks more than his share of the time at the table
30 Actively moves about most of the day

Note—Adapted from Newcomb (1929, p 27)

Newcomb's tables 18, 19, 22, and 23 (1929, pp 118–119, pp 122–123) were selected for re-analysis. These tables consist of the intercorrelations among pairs of the 26 behaviors (Newcomb provides only partial matrices), calculated separately for the daily records and the ratings for each group of boys. The percentage of days an item appeared in a child's behavior was used by Newcomb for calculating intercorrelations among behaviors in the daily records.

In order to obtain judgments of "what conceptually goes with what" in Newcomb's 26 behaviors based on pre-existing rater beliefs, 10 University of Chicago students in one of the author's classes were asked to make paired comparison judgments of conceptual similarity on a 7-point scale (1 = very dissimilar conceptually, 7 = very similar conceptually). Prior to making their conceptual similarity judgments the students were told they would be presented with descriptions about the behavior of boys at a summer camp. Average scores over these 10 sets of judgments for each pair of items were used in all further analyses. The judgments of conceptual similarity were made forty-five years after Newcomb's associates made their ratings.
The re-analysis of the data consists of a comparison of the degree of similarity of the relationships among behaviors in the pre-existing conceptual scheme of the student judges, in the recall-based ratings of Newcomb's six observers, henceforth referred to as "rated behavior," and in the daily records, henceforth referred to as "actual behavior," of two different samples of boys. The degree of similarity of the relationships among behaviors from these three sources is determined by correlating over the intercorrelations of equivalent pairs of behaviors from each pair of the three sources. In other words, three partial matrices of intercorrelation coefficients are themselves compared with each other by correlating over the cells of each pair of matrices. The nonparametric statistic Spearman's rank-order correlation coefficient \( r_s \) was used in the analysis.

Results

The findings of the re-analysis of Newcomb's data appear in Figures 1, 2, and 3. They can be summarized as follows:

1. The relationships among items in rated behavior are more similar to pre-existing conceptual schemes (range of similarity \( r_s = 74-83 \)) than to actual behavior (range of similarity \( r_s = 35-51 \)).

2. The relationships among items in rated behavior remain stable over the two groups of boys \( (r_s = 78) \) in spite of changes in the relationships in actual behavior over the two groups \( (r_s = 41) \).

3. Pre-existing conceptual schemes (notions of "what is like what") are not very similar to the relationships among items in actual behavior (range of similarity \( r_s = 47-51 \)).

4. Rated behavior is almost entirely under the influence of pre-existing conceptual schemes and corresponds to actual behavior relationships only to the extent pre-existing conceptual schemes happen to partially coincide with actual behavior.

In more detail, Figure 1 presents a direct comparison of the degree of similarity of "what goes with what" in actual behavior, and in rated behavior, and "what is like what" in pre-existing conceptual schemes. Newcomb presents correlations based on ratings and based on daily records for the same 64 pairs of the behaviors over both groups of boys. These four sets of correlations are compared in Figure 1 with each other and with the pre-existing conceptual scheme of the student judges. A striking feature of Figure 1 is the degree of similarity of rated behavior.
Figure 1  Comparison of the organization of items in actual behavior, rated behavior, and the pre-existing conceptual scheme

Note—Similarity (Spearman's $r$) of matrices of association among extroverted and introverted items of behavior (Newcomb's list) for two groups of boys based on daily records (actual behavior), recall-based observer ratings (rated behavior), and conceptual similarity judgments (pre-existing conceptual scheme) $N =$ all 64 cells of Newcomb's partial matrices

to the pre-existing conceptual scheme ($r_s = 80$ on the average) and the rather unimpressive degree of similarity of rated behavior to actual behavior ($r_s = 44$ on the average) Moreover, from one group of boys to the next, rated behavior is more stable than actual behavior ($r_s = 78$ in comparison to 41) Perhaps most remarkable of all is the fact that the relationships among items in rated behavior in the first group of boys is no less similar to actual behavior relationships in the second group of boys than
is the rated behavior of the second group of boys, and vice versa \((r_s = 40\) and \(38, 51\) and \(51\))

The actual behavior rated seems to make little difference for the relationships among items revealed in rated behavior. This can be demonstrated in a dramatic way. There is an ambiguity in the results shown in Figure 1. The very modest degree of similarity of rated behavior to actual behavior may reflect a true sensitivity to actual behavior relationships (no matter how attenuated), or it may be the spurious consequence of domination by a conceptual scheme which in this case modestly overlaps with actual behavior. This latter alternative is in fact the case.

The relationships among items in rated behavior replicate pre-existing understandings of “what is like what” even when these understandings directly conflict with actual behavior. This can be seen in Figures 2 and 3. For his second group of boys, Newcomb provides 110 correlations between pairs of items for rated behavior and again for actual behavior. These two sets of correlations were compared with each other and with the pre-existing conceptual scheme of the student judges in two ways. First, the full 110 correlations were compared. These results as presented in Figure 2 closely parallel the results in Figure 1. Rated behavior is very much like the pre-existing conceptual scheme \((r_s = 74)\) and not much like actual behavior \((r_s = 35)\). Actual behavior relationships partially overlap with the pre-existing conceptual scheme \((r_s = 51)\), but we cannot be sure whether the 35 fit of rated behavior and actual behavior is a true sensitivity of judgment to behavior or the artifact of the domination of ratings by a pre-existing conceptual scheme which partially fits actual behavior. This ambiguity was resolved by a second type of analysis. A subset of item correlations was selected where actual behavior relationships were the opposite of relationships in the pre-existing conceptual scheme. This was done by using 33 item correlations in which either the degree of conceptual association of the student judges for the pair was above the mean degree of conceptual association (4.1 on a 7-point scale) and the actual behavior correlation for that pair below the mean actual behavior correlation (10) or the converse.

The results of this type of analysis are dramatic. When actual behavior is opposed to pre-existing understandings of “what is
Figure 2  Comparison of the organization of items in actual behavior, rated behavior, and the pre-existing conceptual scheme when actual behavior overlaps with the pre-existing conceptual scheme $N = \text{all 110 cells of Newcomb's partial matrix}$

Note — Similarity (Spearman's $r_s$) of matrices of association among extroverted and introverted items of behavior (Newcomb's list) for group 2 based on daily records (actual behavior), recall-based observer ratings (rated behavior) and conceptual similarity judgments (pre-existing conceptual scheme)

like what” ($r_s = -36$), rated behavior replicates the conceptual scheme ($r_s = 84$) and conflicts with actual behavior relationships ($r_s = -27$) (see Figure 3) Thus it is clear that recall-based judgments on rating forms merely reproduce pre-existing conceptual associations with little sensitivity to the relationships among items in actual behavior

The data presented so far discriminates among the three interpretations of the finding that rating factors can be independently derived from conceptual criteria There seems to be a causal distorting relationship between conceptual factors and rating factors, i.e., conceptual factors control judgment and are not necessarily related to actual behavior In the light of this evidence, a purely conceptual interpretation is the “preferred-until-shown-otherwise” interpretation of the factors derived from rat-
Figure 3  Comparison of the organization of items in actual behavior, rated behavior, and the pre-existing conceptual scheme when actual behavior conflicts with the pre-existing conceptual scheme. \( N = 33 \) cells of Newcomb's partial matrix where the conceptual similarity score is greater than 4.1 and the actual behavior correlation is less than 0.1 or the conceptual similarity score is less than 4.1 and the actual behavior correlation is greater than 0.1.

Note—Similarity (Spearman's \( r_s \)) of matrices of association among extroverted and introverted items of behavior (Newcomb's list) for group 2 based on daily records (actual behavior), recall-based observer ratings (rated behavior), and conceptual similarity judgments (pre-existing conceptual scheme).

EXPERIMENT II  CONCEPTUAL OVERLAP AND INTERPERSONAL RATINGs  BALEs'S SMALL GROUP SPACE

Method

Bales (1968, p. 467, 1970, p. 30) proposes three factors for describing small-group structure and classifying the personality and role-
types of individuals. The three factors are "power," "likeability," and "task-orientation." They are one of a species of two- and three-factor attempts to classify behavior in groups (see Bales, 1968 for a review). The factors are conceptualized as a property-space with three dimensions.

Bales's three dimensional classification of personality is statistically constructed from the intercorrelations among hundreds of measures on 60 subjects who participated in 12 5-person groups. Most of Bales's measures rely on respondent judgments of the type we have been discussing. These include interpersonal ratings by group members and outside observers, and self-reports on three personality inventories (the MMPI, Cattell's 16 PF test, and Thurstone's temperament test). On the basis of his statistically-constructed classification of personality and group role types, Bales (1970, pp 6–7) has designed a relatively short interpersonal rating form for the purpose of person-classification. The rating form consists of the following 26 items with known loadings in his three dimensional classification:

1. Resentful
2. Admired
3. Dominating
4. Speaks like an autocratic authority
5. Devalues himself
6. Makes others feel he admires them
7. Especially addressed when others have serious opinions about which they want confirmation
8. Stands for the most conservative ideas and beliefs of the group
9. Rejects religious belief generally
10. Liked
11. Feels that others are generally too conforming to conventional social expectations
12. Always tries to speak objectively
13. Thinks of himself as entertaining
14. Warm and personal
15. Accepts failure and withdrawal for himself
16. Identifies himself with the underprivileged
17. Demands pleasure and gratification
18. Believes that equality and humanitarian concerns for others are very important
19. Introverted
20. Believes it is necessary to sacrifice the self for higher values
21. Assumes responsibility for task leadership
Receives a lot of interaction from others
Valuable for a logical task
Feels that his individual independence is very important
Personally involved in the group
Passively withholds cooperation

Note—Adapted from Bales (1970, pp 6-7)

It has been shown previously (Shweder, 1972, p 24) that the rated behavior relationships among the items on Bales's interpersonal rating form are essentially consistent with Bales's three dimensional classification. The concern here is to demonstrate that Bales's three dimensional classification as revealed in rated behavior is, and in the light of our previous analysis may only be, a pre-existing conceptual scheme (a theory about "what is like what") in the minds of raters.

Table 1 compares the statistical classification of Bales's 26 items derived from two sources, rated behavior and pre-existing understandings of "what is like what." Rating dimensions were arrived at as follows: Twenty-five undergraduates at Harvard University, participants in a single self-analytic group for over 40 hours over a three-month period, rated all the members of their group on Bales's 26 interpersonal behavior items. The ratings were dichotomous, yes or no. R. Freed, Bales, and Stephen P. Cohen were kind enough to provide this data for analysis. Correlation coefficients (Pearson $r$) were calculated among all possible pairs of items over 600 ratings. They were then scaled in three dimensional space using the multidimensional scaling program (MDSCALE) devised by Kruskal and Shepard (Kruskal & Ling, 1967, Shepard, 1962a, b, 1963) and rotated to a representation of Bales's classification of these items.

In order to obtain judgments of "what conceptually goes with what" based on pre-existing rater beliefs, Bales's 26 items were placed on cards and presented to each of 50 undergraduate students at Harvard University. These students were not participants in the group. The students sorted the items into piles according to the following written instructions:

On the cards before you are written brief descriptions of persons' behavior in a group. Cluster the cards together in terms of the similarity of meaning of the descriptions. That is, make as many piles as you wish, with as many cards in each pile as you wish. Each pile should contain descriptions similar in meaning. Take all the time you need. Feel free to rearrange the cards until you find a set of piles that most satisfies you as similar in meaning. Piles with only one member are not permitted.
Individual difference theory of personality

Table 1  Bales's three dimensional group space Conceptual and rating dimensions compared  MDSCALE spatial coordinates

<table>
<thead>
<tr>
<th>Item #</th>
<th>Conceptual dimension</th>
<th>Rating dimension</th>
<th>Conceptual dimension</th>
<th>Rating dimension</th>
<th>Conceptual dimension</th>
<th>Rating dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.38</td>
<td>0.01</td>
<td>-0.89</td>
<td>-0.99</td>
<td>-0.12</td>
<td>-0.44</td>
</tr>
<tr>
<td>2</td>
<td>0.18</td>
<td>0.34</td>
<td>0.74</td>
<td>0.71</td>
<td>0.27</td>
<td>0.30</td>
</tr>
<tr>
<td>3</td>
<td>0.71</td>
<td>0.77</td>
<td>-0.82</td>
<td>-0.54</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td>4</td>
<td>0.63</td>
<td>0.65</td>
<td>-1.00</td>
<td>-0.71</td>
<td>-0.05</td>
<td>0.21</td>
</tr>
<tr>
<td>5</td>
<td>-1.13</td>
<td>-0.99</td>
<td>-0.06</td>
<td>0.12</td>
<td>-0.18</td>
<td>-0.43</td>
</tr>
<tr>
<td>6</td>
<td>-0.18</td>
<td>-0.15</td>
<td>0.89</td>
<td>1.06</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.62</td>
<td>0.66</td>
<td>0.41</td>
<td>0.31</td>
<td>0.49</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>0.43</td>
<td>-0.34</td>
<td>-0.68</td>
<td>-0.78</td>
<td>0.29</td>
<td>0.67</td>
</tr>
<tr>
<td>9</td>
<td>-0.09</td>
<td>0.03</td>
<td>-0.44</td>
<td>-0.12</td>
<td>0.30</td>
<td>-0.96</td>
</tr>
<tr>
<td>10</td>
<td>0.65</td>
<td>-0.01</td>
<td>0.89</td>
<td>1.00</td>
<td>-0.27</td>
<td>0.20</td>
</tr>
<tr>
<td>11</td>
<td>-0.23</td>
<td>0.16</td>
<td>-0.53</td>
<td>-0.53</td>
<td>-0.66</td>
<td>-0.73</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.21</td>
<td>-0.02</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>13</td>
<td>0.65</td>
<td>0.64</td>
<td>0.00</td>
<td>-0.13</td>
<td>-0.46</td>
<td>-0.65</td>
</tr>
<tr>
<td>14</td>
<td>0.38</td>
<td>-0.10</td>
<td>1.04</td>
<td>1.03</td>
<td>-0.36</td>
<td>-0.18</td>
</tr>
<tr>
<td>15</td>
<td>-1.22</td>
<td>-1.14</td>
<td>-0.09</td>
<td>-0.46</td>
<td>0.16</td>
<td>-0.48</td>
</tr>
<tr>
<td>16</td>
<td>-0.71</td>
<td>-0.72</td>
<td>-0.22</td>
<td>0.36</td>
<td>-0.71</td>
<td>-0.41</td>
</tr>
<tr>
<td>17</td>
<td>0.45</td>
<td>0.51</td>
<td>-0.55</td>
<td>-0.78</td>
<td>-0.48</td>
<td>-0.43</td>
</tr>
<tr>
<td>18</td>
<td>-0.29</td>
<td>-0.41</td>
<td>0.61</td>
<td>0.80</td>
<td>-0.38</td>
<td>-0.01</td>
</tr>
<tr>
<td>19</td>
<td>-1.12</td>
<td>-1.28</td>
<td>-0.26</td>
<td>-0.15</td>
<td>0.38</td>
<td>0.17</td>
</tr>
<tr>
<td>20</td>
<td>-0.70</td>
<td>-0.46</td>
<td>-0.09</td>
<td>0.42</td>
<td>-0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>21</td>
<td>0.79</td>
<td>0.75</td>
<td>0.06</td>
<td>0.02</td>
<td>0.20</td>
<td>0.44</td>
</tr>
<tr>
<td>22</td>
<td>0.73</td>
<td>0.79</td>
<td>0.85</td>
<td>-0.03</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>23</td>
<td>0.28</td>
<td>0.47</td>
<td>0.45</td>
<td>0.42</td>
<td>0.92</td>
<td>0.66</td>
</tr>
<tr>
<td>24</td>
<td>-0.05</td>
<td>0.34</td>
<td>-0.83</td>
<td>-0.75</td>
<td>0.18</td>
<td>0.24</td>
</tr>
<tr>
<td>25</td>
<td>0.34</td>
<td>0.40</td>
<td>1.11</td>
<td>0.64</td>
<td>0.00</td>
<td>-0.05</td>
</tr>
<tr>
<td>26</td>
<td>-0.76</td>
<td>-0.90</td>
<td>-0.78</td>
<td>-0.90</td>
<td>-0.19</td>
<td>-0.19</td>
</tr>
</tbody>
</table>

Approximate r, of conceptual dimension and rating dimension  86 .88  70

An association measure among all pairs of items was calculated from the sorting task data. The degree of association between any two items was based on three sources of information: the number of informants placing the two items in the same pile, the number of other items in the pile shared by the two items, and the total number of piles made by the informant (Burton, 1968, pp. 81-83 discusses this measure of association). The conceptual association matrix was scaled in three dimensional space using the MDSCALE program and rotated to Bales's classification of the items.

Results

Table 1 lists the three dimensional spatial coordinates of the 26 items in both rated behavior and the shared pre-existing conceptual scheme. The two classifications are extremely similar, i.e., each of the three conceptual dimensions is similar to its cor-
responding rating dimension. The rank-order correlation \( r_s \) of the 26 items on the "power" dimension is approximately 86 between the conceptual dimension and the rating dimension. For the "likeability" dimension, it is 88, for the "task-orientation" dimension, it is 70. In the light of our earlier finding that respondent judgments are entirely under the control of pre-existing conceptual schemes which may themselves have a relatively poor fit to actual behavior, Bales's classification of small group behavior seems somewhat equivocal.

**EXPERIMENT III  CONCEPTUAL OVERLAP AND RETROSPECTIVE QUESTIONNAIRES  SEARS, MACCOBY AND LEVIN'S MOTHER INTERVIEW**

**Method**

Sears, Maccoby and Levin (1957, pp 471–477, pp 516–518) proposed a number of underlying personality traits which they believe influence a mother's child-rearing practices. The underlying traits were statistically constructed applying a factor analysis to the intercorrelations among 44 "theoretically significant" scales which Sears, Maccoby and Levin used to assess the responses of 379 mothers to their well-known mother interview. The two most statistically important factors in the analysis were maternal "permissiveness" and maternal "self-confidence" (family adjustment). The 44 scales follow:

1. Mother's responsiveness to infant's crying
2. Amount of mother's affectionate interaction with baby
3. Amount of fun taking care of small baby
4. Severity of weaning
5. Scheduling of feeding
6. Severity of handling feeding problems
7. Level of demands for table manners
8. Amount of praise for good behavior at the table
9. Severity of toilet training
10. Permissiveness for going without clothes indoors
11. Masturbation permissiveness
12. Permissiveness concerning sex play (among children)
13. Standards for neatness and orderliness
14. Restrictions relating to care of house and furniture
15. Strictness about bedtime
16. Strictness about noise
17. Restrictions on physical mobility
18. Mother's realistic standards for child's obedience
19 Amount of praise for obedience
20 Mother's response to dependency
21 Amount of mother's affectionate demonstrativeness
22 Does mother find time to play with child
23 How important that child do well in school
24 Permissiveness for aggression toward siblings
25 Praise for playing nicely with children
26 Demands for child to be aggressive
27 Permissiveness for inappropriate aggression toward other children
28 Permissiveness for aggression toward parents
29 Punishment for aggression toward parents
30 Extent of use of tangible rewards
31 Extent of use of physical punishment
32 Extent of use of deprivation of privileges
33 Extent of use of reasoning
34 How often threaten to punish, then not follow through
35 Affectional bond, father to child
36 Which parent disciplines
37 Parents' agreement on child-rearing policies
38 Responsibility for child-rearing policies
39 Mother's evaluation of father
40 Mother's reaction to pregnancy
41 Husband's reaction to wife's pregnancy
42 Mother's dissatisfaction with current situation
43 Mother's self-esteem
44 Mother's child-rearing anxiety

Note—From Sears, Maccoby and Levin (1957, pp 516–518)

Sears, Maccoby and Levin's two major purported underlying personality traits of mothers are pre-existing conceptual factors in the minds of the mothers who respond to their interview, i.e., they can be reproduced by having a small number of informants simply judge the conceptual similarity of all pairs of 44 scales. On the basis of the previous discussion of the conceptual distortion of respondent judgments, it seems appropriate to doubt that these underlying traits of maternal personality are anything more than ways in which mothers classify themselves. There certainly is little reason to believe they are factors influencing maternal child-rearing behavior. An explanation follows.

In order to obtain judgments of "what conceptually goes with what" based on pre-existing rater beliefs, Sears, Maccoby and Levin's 44 scales were converted into declarative assertions (e.g., "strictness
**Table 2** Sears, Maccoby and Levin's maternal personality factors conceptual and interview factors compared  
Factor loadings

<table>
<thead>
<tr>
<th>Table 6 Scale #</th>
<th>Conceptual factor</th>
<th>Interview factor</th>
<th>Conceptual factor</th>
<th>Interview factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-21</td>
<td>-15</td>
<td>.89</td>
<td>05</td>
</tr>
<tr>
<td>2</td>
<td>-15</td>
<td>02</td>
<td>.93</td>
<td>03</td>
</tr>
<tr>
<td>3</td>
<td>-13</td>
<td>-06</td>
<td>.93</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>11</td>
<td>-29</td>
<td>-14</td>
</tr>
<tr>
<td>5</td>
<td>88</td>
<td>09</td>
<td>-36</td>
<td>-11</td>
</tr>
<tr>
<td>6</td>
<td>88</td>
<td>22</td>
<td>-35</td>
<td>-06</td>
</tr>
<tr>
<td>7</td>
<td>95</td>
<td>37</td>
<td>05</td>
<td>-02</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>-07</td>
<td>62</td>
<td>02</td>
</tr>
<tr>
<td>9</td>
<td>95</td>
<td>34</td>
<td>-15</td>
<td>-25</td>
</tr>
<tr>
<td>10</td>
<td>-77</td>
<td>-66</td>
<td>34</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>-79</td>
<td>-70</td>
<td>31</td>
<td>04</td>
</tr>
<tr>
<td>12</td>
<td>-82</td>
<td>-71</td>
<td>19</td>
<td>06</td>
</tr>
<tr>
<td>13</td>
<td>90</td>
<td>.35</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>14</td>
<td>89</td>
<td>48</td>
<td>22</td>
<td>-01</td>
</tr>
<tr>
<td>15</td>
<td>90</td>
<td>10</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td>93</td>
<td>38</td>
<td>03</td>
<td>-13</td>
</tr>
<tr>
<td>17</td>
<td>85</td>
<td>18</td>
<td>-22</td>
<td>08</td>
</tr>
<tr>
<td>18</td>
<td>-01</td>
<td>32</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>49</td>
<td>05</td>
<td>77</td>
<td>-15</td>
</tr>
<tr>
<td>20</td>
<td>33</td>
<td>38</td>
<td>-74</td>
<td>-07</td>
</tr>
<tr>
<td>21</td>
<td>-15</td>
<td>-03</td>
<td>92</td>
<td>28</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>-04</td>
<td>-80</td>
<td>-24</td>
</tr>
<tr>
<td>23</td>
<td>62</td>
<td>32</td>
<td>59</td>
<td>-02</td>
</tr>
<tr>
<td>24</td>
<td>-63</td>
<td>-31</td>
<td>12</td>
<td>-12</td>
</tr>
<tr>
<td>25</td>
<td>38</td>
<td>-03</td>
<td>75</td>
<td>-01</td>
</tr>
<tr>
<td>26</td>
<td>-40</td>
<td>-23</td>
<td>06</td>
<td>-05</td>
</tr>
<tr>
<td>27</td>
<td>-68</td>
<td>-30</td>
<td>-43</td>
<td>-09</td>
</tr>
<tr>
<td>28</td>
<td>-80</td>
<td>-.51</td>
<td>-17</td>
<td>-09</td>
</tr>
<tr>
<td>29</td>
<td>78</td>
<td>43</td>
<td>03</td>
<td>-13</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>04</td>
<td>73</td>
<td>14</td>
</tr>
<tr>
<td>31</td>
<td>25</td>
<td>31</td>
<td>-67</td>
<td>-15</td>
</tr>
<tr>
<td>32</td>
<td>.56</td>
<td>16</td>
<td>13</td>
<td>08</td>
</tr>
<tr>
<td>33</td>
<td>09</td>
<td>-04</td>
<td>89</td>
<td>07</td>
</tr>
<tr>
<td>34</td>
<td>-32</td>
<td>06</td>
<td>-30</td>
<td>-35</td>
</tr>
<tr>
<td>35</td>
<td>09</td>
<td>-02</td>
<td>-75</td>
<td>-48</td>
</tr>
<tr>
<td>36</td>
<td>62</td>
<td>15</td>
<td>-13</td>
<td>01</td>
</tr>
<tr>
<td>37</td>
<td>05</td>
<td>-01</td>
<td>-50</td>
<td>-60</td>
</tr>
<tr>
<td>38</td>
<td>-38</td>
<td>-20</td>
<td>-37</td>
<td>00</td>
</tr>
<tr>
<td>39</td>
<td>-11</td>
<td>-06</td>
<td>80</td>
<td>.56</td>
</tr>
<tr>
<td>40</td>
<td>03</td>
<td>-07</td>
<td>-.89</td>
<td>-42</td>
</tr>
<tr>
<td>41</td>
<td>-04</td>
<td>-13</td>
<td>-.82</td>
<td>-51</td>
</tr>
<tr>
<td>42</td>
<td>-04</td>
<td>-08</td>
<td>-93</td>
<td>-47</td>
</tr>
<tr>
<td>43</td>
<td>-20</td>
<td>-10</td>
<td>85</td>
<td>.59</td>
</tr>
<tr>
<td>44</td>
<td>21</td>
<td>10</td>
<td>-70</td>
<td>-.56</td>
</tr>
</tbody>
</table>

*Approximate r, of conceptual and interview factor*  
89 79

*Note—Loadings on interview factor are from Sears, Maccoby and Levin (1957, p 516-518)*
about bedtime" became "Mother is strict about bedtime") and presented to five University of Chicago graduate students. The students were asked to make paired comparison judgments of conceptual similarity on a 7-point scale. Prior to making their conceptual similarity judgments, the students were told they would be presented with descriptions of the behaviors and attitudes of mothers toward their children. The student judges participated in none of the other tasks reported in this article. Average scores over these five sets of judgments for each pair of items were used in all further analyses.

The pre-existing conceptual scheme shared by these five students was analyzed in the following way. Their pooled and averaged conceptual similarity judgments were placed in a full $44 \times 44$ matrix. Intercorrelations (Pearson $r$) among the 44 scales were calculated by treating the rows as subjects and the columns as variables. The resulting matrix of intercorrelations was factor analyzed using a principal axes analysis and rotated using a varimax rotation. Seven factors were extracted, paralleling Sears, Maccoby, and Levin's analysis. The first two factors were of equal importance and accounted for 70 percent of the variance in the data. A third factor accounted for only 7 percent of the variance and was not considered.

Results

Table 2 lists the factor loadings of the 44 scales on the first two factors in both the maternal interviews of 379 mothers and the pre-existing conceptual scheme of five students. Both pre-existing conceptual factors are very similar to their corresponding interview factors. The rank-order correlation ($r_o$) of the 44 scales on the "permissiveness" factor is 89 between the conceptual factor and the interview factor. For "self-confidence" (family adjustment), it is 79. Sears, Maccoby and Levin's third factor, maternal "warmth," could not be recovered from the conceptual similarity judgments of the student judges. Perhaps if more than just five students had made conceptual similarity judgments, it may have been extractible, or perhaps like many of the other factors in their analysis, it was a minor (and one would guess unstable) statistical phenomenon. In any case, the ease with which Sears, Maccoby, and Levin's two most important factors can be independently derived from the conceptual associations of only five students both demonstrates they are pre-existing conceptual factors, and suggests the equivocality of their status at maternal personality traits influencing child rearing practices.
EXPERIMENT IV CONCEPTUAL OVERLAP AND PERSONALITY INVENTORIES THE MMPI

Method

Self-reports on personality inventories may be subject to the same type of conceptually determined bias hypothesized for interpersonal ratings and interview accounts of behavior. Self-reported relationships among items on the MMPI seem to be derivable from purely conceptual criteria.

The MMPI is a self-report instrument originally designed and standardized with an eye towards the assessment of psychopathology. Its 566 items lean heavily in the direction of psychiatric diagnostic considerations. Both the size and the nature of the MMPI item pool place limitations on the scope of possible conceptual/self-report comparisons. First, it is not feasible to ask students to judge the degree of conceptual similarity among 566 items. Second, the distribution of responses in a psychologically normal population to many of the questions makes an analysis of item association impossible. For example, “evil spirits possess me at times” (MMPI item #27) is answered in the affirmative by one percent of college males (Dahlstrom and Welsh, 1960, p 418). Thus, the following analysis is presented as a partial, tentative, and at best preliminary report on the influence of pre-existing conceptual schemes on inventory self-reports. The MMPI protocols discussed below were collected by Dr. Bertram J. Cohler from 172 female Peace Corps volunteers. They will be the subject of a forthcoming report on conceptual bias in the MMPI. A single example from an initial investigation of this material seems appropriate in this context.

The following are six MMPI items whose distribution of yes-no responses permits a calculation of item correlations in a normal population:

1. My feelings are not easily hurt (MMPI #79)
2. I do not mind being made fun of (MMPI #91)
3. I like to go to parties and other affairs where there is lots of loud fun (MMPI #99)
4. I wish I were not so shy (MMPI #201)
5. I tend to be on my guard with people who are somewhat more friendly than I had expected (MMPI #348)
6. I have several times given up doing a thing because I thought too little of my ability (MMPI #357)

The following analysis was performed in order to obtain ratings of “what conceptually goes with what” based on pre-existing subject
Individual difference theory of personality

Table 3 Conceptual similarity and item endorsement on the MMPI, degree of conceptual similarity (7.0 = conceptually very similar) and item intercorrelation (Pearson r, N = 173) for six MMPI items

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.8 (24)</td>
<td>5.8 (00)</td>
<td>2.0 (-29)</td>
<td>2.8 (-08)</td>
<td>2.2 (-12)</td>
</tr>
<tr>
<td>2</td>
<td>5.8 (14)</td>
<td>2.4 (-16)</td>
<td>2.2 (-08)</td>
<td>2.4 (-11)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.2 (09)</td>
<td>2.2 (-09)</td>
<td>3.4 (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.6 (10)</td>
<td>5.8 (17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note—Approximate r of conceptual similarity matrix and self report item intercorrelation matrix = 70

beliefs, five University of Chicago graduate students were asked to make paired comparison judgments of conceptual similarity on a 7-point scale (7 = conceptually very similar). Prior to making their conceptual similarity judgments the students were told they would be presented with a person's description of himself. The student judges participated in none of the other tasks reported in this article. Average scores over these five sets of judgments were used in the analysis. They are presented in matrix form in Table 3. The intercorrelations (Pearson r) over the MMPI protocols of 172 female Peace Corps volunteers were also calculated for all pairs of items. These are also presented in Table 3.

Results

The relationships among the items in the pre-existing conceptual scheme of the five student judges is similar to the self-reported relationships among the items in the MMPI protocols of 172 Peace Corps women (r = 70 approximately). The same level of conceptual likeness and self-report item relationship similarity (r = 70) for selected items from Drake's "Social Introversion" scale of the MMPI (as specified in Dahlstrom and Welsh, 1960) has also been discovered. It would be presumptuous on the basis of so partial and preliminary an analysis to enter into the controversy over the interpretation of the first two purported personality factors (sometimes called "ego-resilience" and "ego-control") (Block, 1965) extracted from the MMPI scales (themselves composed of MMPI items) (See e.g., Block, 1965, Edwards, 1953, 1961, Edwards and Diere, 1962, Jackson and Messick, 1961, Wiggins, 1966). However, the above analysis strongly suggests a methodological interpretation of these two
factors An increase in the degree of conceptual similarity of two items increases the probability that they will both be endorsed or denied by the same person. Thus, the pattern of relationships among MMPI scales should reflect the average degree of conceptual likeness among scale items.

**DISCUSSION**

The existence of highly interpretable item intercorrelations on interpersonal checklists, personality inventories and questionnaire interviews has supported a general impression of the relevancy of the concept “personality” in its widely acknowledged “individual difference” sense. The validity of this corpus of supportive data for the “personality” concept now seems to be questionable. Evidence has been presented that such personality assessment procedures offer a temptation, typically succumbed to by respondents, to apply a theory of conceptual likenesses, which is of course highly interpretable since it is about conceptual likenesses, as if it were a theory of behavioral happenings, which it is not. Items alike in concept are judged to be characteristic of the same person even when the conceptual relationships among items do not correspond to the relationships among items in actual behavior.

The nature of the logical error involved in such a substitution of judgments about conceptual likeness for judgments about co-occurring instances can be clarified with a nonbehavioral example. Things can be and typically are alike without co-occurring “Blue” and “green” for example are both “colors.” They are attributes conceptually derivative on the attribute “colored.” Thus, any instance of “blue” is always an instance of “color” but not the converse. The same is true of “green.” But one could not properly reason that if “blue” and “green” are related just so, as conceptual derivatives of “color,” then they probably will co-occur as attributes of “colored things.” It might be that no “colored things” were both “blue” and “green” and we would still feel a conceptual likeness between “blue” and “green” as “colors.” Things can be alike without co-occurring.

The same independence of conceptual likeness and co-occurrence likelihood should exist for items of behavior. “Seeks reassurance” and “seeks help” may refer to attributes conceptually derivative on the attribute “dependent” (desiring to be cared for)
Individual difference theory of personality

If "seeking reassurance" and "seeking help" were never to occur together in the same person or same behavioral context or whatever, their conceptual likeness as attributes would still remain. Yet items similar in concept (e.g., "seeks reassurance" and "seeks help") are typically judged to be characteristic of the same person on verbal report personality assessment procedures, thereby creating the "illusion" that personality factors are responsible for the consistent display of behaviors expressive of the factor.

The mechanism by which "illusory" relationships are produced by pre-existing conceptual associations seems to be a rather simple one and is related to a characteristic feature of the personality assessment task on checklists, inventories and questionnaires. They are a kind of situation that encourages a "rush to judgment" even when, as is typical, the respondent does not have access to relevant information or the relevant information is unavailable, i.e., the respondent either knows too much or too little about the person judged. Random guessing under conditions of ignorance would, of course, reduce the size of correlations among items and not alter the pattern of interitem relationships, but guesses are never random when the judge has a pre-existing understanding of "what is like what." Random error becomes systematic error by consistently erring in the direction of the pre-existing conceptual scheme, the result of which is to increase the size of interitem correlations while organizing it into an erroneous but coherent pattern. Table 4 provides supportive evidence for this claim.

Table 4 displays the differential magnitude of rated behavior correlations and actual behavior correlations among pairs of items in Newcomb's 26-item behavior list in relation to their degree of conceptual likeness. In Table 4 item pairs in the quintile where pre-existing understandings that items are alike are the strongest (mean conceptual similarity score of 6.2 on a 7-point scale) have an average rated behavior correlation 53 and an average actual behavior correlation 29. Item pairs in the quintile where pre-existing understandings that items are unlike are the strongest (mean conceptual similarity score of 2.1 on a 7-point scale) have an average rated behavior correlation -50 and an average actual behavior correlation -10. One final feature of the data in Table 4 is worth noting. The understanding that items are alike seems to have a greater influence on judgment than the understanding
Table 4  The differential magnitude of rated behavior correlations and actual behavior correlations The differential influence of pre-existing conceptual schemes on rated behavior correlations

<table>
<thead>
<tr>
<th>N</th>
<th>Conceptual similarity scores (7-point scale)</th>
<th>Rated behavior correlations (correlation points)</th>
<th>Actual behavior correlations (correlation points)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>22</td>
<td>62</td>
<td>29</td>
<td>53</td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>22</td>
<td>40</td>
<td>27</td>
<td>-03</td>
</tr>
<tr>
<td>22</td>
<td>29</td>
<td>31</td>
<td>-27</td>
</tr>
<tr>
<td>22</td>
<td>21</td>
<td>20</td>
<td>-50</td>
</tr>
</tbody>
</table>

Note—Means and standard deviations for group 2 rated behavior correlations, actual behavior correlations, and conceptual similarity scores, divided into quintiles in terms of the degree of conceptual similarity for pairs of extroverted and introverted items in Newcomb's 26-item behavior list

that they are unalike. There is less dispersion of rated behavior correlations (standard deviation = 28 correlation points) in the quintile with the highest average conceptual similarity score than in the quintile of rated behavior correlations (standard deviation = 43 correlation points) with the lowest average conceptual similarity score.

Respondents systematically err in trying faithfully to answer the questions put to them on checklists, inventories, and questionnaires, but no more so it would seem than personality theorists. That is to say, the special way items of behavior are assumed to be causally related to hypothesized underlying variables in the factor-analytic model has also deceptively “confirmed” the applicability of the concept “personality” in its “individual difference” sense. The causal assumption is that a set of behaviors are related to one another only as effects of the same cause, the underlying variable (trait, disposition, factor, dimension). None of the behaviors is assumed to be a direct cause or effect of another. The predicted positive intercorrelations among the behaviors are viewed as spurious consequences of the causal influence (trait, disposition, factor, dimension) reducible to zero if the underlying variable is controlled (Blalock, 1964, pp 167–169). This kind of relatedness of behaviors as effects of a common hypothesized cause is similar to the kind of relatedness characteristic of a set of attributes conceptually derivative on the same prior attribute. If such conceptually derivative attri-
Individual difference theory of personality

Attributes are responded to in like fashion on checklists, inventories, and questionnaires, they will falsely "appear" to co-occur in a way consistent with trait-factor assumptions. This "appearance" as we have seen is "illusory." Without an appeal to actual performance standards we have no grounds for claiming that the traits, dispositions, underlying states, scales, factors and dimensions discovered in interpersonal checklists, personality inventories and questionnaire interviews are anything more than stimulus (in contrast to person) factors, i.e., ways people classify things that are conceptually alike. Such classificatory factors may be relevant in the formation, manipulation, and perpetuation of "social impressions" yet have no wider domain of relevancy.

To so explicitly separate, as has just been done, the structure of "social impressions" from the structure of behavioral instances is to directly broach a question which may have been on the reader's mind for some time. How can our schemes for classifying people, our conceptual dimensions for describing one another, fail to be more or less accurate inductive generalizations about behavior and at the same time continue to be part of our linguistic and cultural inheritance? How is it that a conceptual scheme that is demonstrably not a codification of accumulated behavioral experience persists uncorrected?

The problem of how to interpret the discrepancy between mental associations and behavioral connections is an important one for future research, and it would be folly to try to solve it at this time. However, it does seem appropriate to speculate about two possible tacks with regard to a solution. On the one hand to pose the question (how can a mistaken conceptual scheme maintain itself in the face of disconfirming behavioral experience?) may be to miss the point. It is possible that the function of person classification in everyday affairs is other than drawing detailed inferences about behavioral outcomes, e.g., it may function to do no more than create types of persons and in the very act of type creation assure us that people are "of a kind" (have stable natures). When the process of social interaction is closely observed, the implicit understandings of "what actors are" (understandings required to successfully interact at all) may have more to do with the actor's momentary, fleeting interpretation.
of immediately preceding acts in interaction sequences (see Raush, 1965) than global personality attributes.

On the other hand opportunities to disconform conceptual associations may be limited by both cognitive and contextual constraints. Conceptual associations may be described as propositions about what "sensibly" goes together. When they are about items of behavior they define the very nature of "consistent" behavior for those who share them, i.e., from the perspective of those sharing the conceptual associations behavior is "orderly" to the extent it can be said to parallel these associations. The intellectual tendency to confirm what is "sensible" and dismiss (forget, reinterpret, explain away or fail to retrieve) what "lacks sense" must be overwhelming in everyday affairs. Wason and Johnson-Laird (1972) have in fact elaborately documented the bias towards redundant verification (and away from disconfirmation) in hypothesis testing, and Garner (1966) has discussed the tendency for certain cognitive processes to drift in the direction of what is informationally easy to specify (in his terms "good forms", in my terms "cultural sense"). Finally, behavioral opportunities to disconfirm conceptual associations are probably minimal. Most of our fellows are seen in certain restricted contexts where (1) only an attenuated variety of behaviors (well below the diversity required to "test" our conceptual associations) ever occur, and (2) because of the stability of the behavioral setting in relationship to those fellows we do frequently observe, those behaviors that do occur are decent candidates for true longitudinal consistency. Also see Jones and Nisbett (1971) and Tversky and Kahneman (1974) for a more complete discussion of the informational, linguistic and cognitive conditions which tend to contribute "illusory" support for an "individual difference" theory of personality.

The challenge held out to personality psychology by this paper is to empirically support with observational data the applicability of the concept "personality" in its "individual difference" sense. The analyses suggest it is considerably less relevant for accounts of behavior than is generally supposed although the possibility that it is pertinent in certain restricted domains (e.g., with regard to hereditable characteristics such as activity level), should not be dismissed.

Exemplary of what is predicted about comparisons between
observational data and data from verbal report assessment procedures is the discrepancy between the findings of Beller (1957) and Sears (1963) on dependency motivation. Beller uses observer ratings of children's social behavior and discovers high item correlations among ostensible (conceptually alike) indicators of dependency. His data supports the notion that dependency is an underlying motive for behavior. Sears uses immediate scorings of children's social behavior and discovers very low item intercorrelations among a very similar set of ostensible (conceptually alike) indicators of dependency. Little support for an underlying dependency motive is found in his data. Further illuminating evidence on the lack of underlying consistencies in the way people differ can be found in Endler and Hunt (1968), Johnson-Laird, Legrenzi and Legrenzi (1972), Mischel (1968, 1973), Raush, Dittmann and Taylor (1959), Raush, Farbman and Llewellyn (1960), Shweder (1971, 1973), Wason (1969), and Wason and Shapiro (1971).

Curiously, much of what has been said here Newcomb was aware of in 1929. He demonstrated that most of the boys in his summer camp showed inconsistent behaviors in most situations (1929, p. 40). He presented evidence that randomly selected pairs of behaviors were as highly associated in actual behavior as pairs of behaviors from the same purported "trait" (1929, p. 44). He showed that ratings tended to confirm relationships predicted by the introversion vs. extroversion hypothesis which could not be confirmed in actual behavior (1929, p. 111), and he concluded that "these relationships are to be found in the minds of the raters to a greater extent than in the boys' actual behavior" (1929, p. 111).

Summary

Doubts are raised as to the extent of "applicability" of the conceptualization that "personality" consists of more or less stable internal factors that make one person's behavior consistent from one time to another, and different from the behavior other people would manifest in comparable situations. This is done by demonstrating the questionable validity of much of the extant empirical support for the "personality" concept. Respondents on interpersonal checklists, personality inventories and questionnaire interviews are shown to unwittingly substitute a theory
of conceptual likenesses for a description of behavioral co-occurrences. Considerations about similarity are confounded with judgments about probability to such an extent that items alike in concept are inferred to be behaviorally characteristic of the same person even when, as is typically the case, conceptual relationships among items do not correspond to the actual behavioral relationships among items. Examined are extant "personality theories" having to do with children’s social behavior, adult behavior in small groups, maternal socialization practices, and psychopathology. These "theories" are shown to be no more than statements about how respondents (and psychologists) classify things as alike in meaning.

References

Burton, R. Validity of retrospective reports assessed by the multitrait-multimethod analysis. Developmental Psychology, 1970, 3, 1-15
Campbell, D T., & Fiske, D W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105
Edwards, A L The relationship between the judged desirability of a trait and the probability that the trait will be endorsed Journal of Applied Psychology, 1953, 37, 90-93
Edwards, A L Social desirability or acquiescence in the MMPI? A case study with the SD scale Journal of Abnormal Psychology, 1961, 63, 351-359
Edwards, A L & Diers, C Social desirability and the factorial interpretation of the MMPI Educational and psychological measurement, 1962, 22, 501-509
Fiske, D W The limits for the conventional science of personality Journal of Personality, 1974, 42, 1-11
Garner, W R To perceive is to know American Psychologist, 1966, 21, 11-19
Jackson, D & Messick, S Acquiescence and desirability as response determinants on the MMPI Educational and Psychological Measurement, 1961, 21, 771-790
Jones, E E & Nisbett, R E The actor and the observer Divergent perceptions of the causes of behavior In Jones, E E, et al (Eds.), Attribution Perceiving the causes of behavior New York General Learning Press, 1971
Kruskal, J B & Ling, R How to use the Yale version of MDScale, a multidimensional scaling program Unpublished manuscript, 1967
Mischel, W Towards a cognitive social learning reconceptualization of personality Psychological Review, 1973, 80, 252-283
Newcomb, T M The consistency of certain extrovert-introvert behavior patterns in 51 problem boys Teachers College, Columbia University, Contributions to Education, 1929, No 382
Raush, H L, Dittmann, A T, & Taylor, T J Person, setting and change in social interaction Human Relations, 1959, 12, 361-377
Raush, H L, Farbman, I & Llewellyn, L Person, setting and change in social interaction II A normal-control study Human Relations, 1963, 13, 305-332
Sears, R R Dependency motivation In M R Jones (Ed.), Nebraska Symposium on Motivation Lincoln University of Nebraska Press, 1963
Shepard, R The analysis of proximities Multidimensional scaling with an unknown distance function I Psychometrika, 1962a, 27, 125-140
Shepard, R The analysis of proximities Multidimensional scaling with an unknown distance function II Psychometrika, 1962b, 27, 219-246
Shepard, R Analysis of proximities as a technique for the study of information processing in man Human Factors, 1963, 5, 33-48
Shweder, R A Is a culture a situation? Honorable Mention paper 1971 Stirling Award in Culture and Personality Studies sponsored by the American Anthropological Association 1971
Shweder, R A Semantic structures and personality assessment Doctoral dissertation, Harvard University University Microfilms, Ann Arbor, Michigan, Order No 72-79, 584 1972
Shweder, R A The between and within of cross-cultural research Ethos, 1973, 1, 531–543

Tversky, A & Kahneman, D Judgment under uncertainty Heuristics and Biases Science, 1974, 185, 1124–1131


Wason, P C & Shapiro, D Natural and contrived experience in a reasoning problem Quarterly Journal of Experimental Psychology, 1971, 23, 63–71


Wiggins, J Substantive dimensions of self-report in the MMPI item pool Psychological Monographs 80 (whole number 630) 1966


Manuscript received August 5, 1974
This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.