Further Study of Cognitive Processing Models for Inventory Response

Norman Cliff
University of Southern California

An attempt was made to validate for sentence type items a mathematical model for inventory response, which had previously been found to account quite well for adjective checklist responses. Data were gathered from subjects responding under candid and under faking sets. In the former case only limited support for the model was found, but in the latter it seemed highly relevant.

The items that make up personality inventories are by definition verbal statements, which vary in length from a single word to a compound sentence. The respondent indicates the degree to which the statement is true of him. A plausible intuitive model for inventory response is that the respondent evaluates or interprets the content of the item according to the internalized schema he has for such information, compares this interpretation to his self-image or to the *persona* he wishes to present, and transforms the comparison onto the inventory's response scale. How to translate this intuitive model into an empirically testable form is a scientific problem.

One general strategy is to use some method such as multidimensional scaling (MDS) to define the internalized semantic schema and then see if the degree of endorsement of an item is systematically related to where it is located in the schema. Using MDS, each item j is assigned to a set of coordinates v_{jm} on each of the dimensions m. Then x_{ij}, the degree of endorsement of item j by person i, is studied as a function of the v_{jm}:

$$x_{ij} = f(v_{j1}, v_{j2} \cdots),$$

and an attempt may be made to specify the function.

Cliff, Bradley, and Girard (1973) employed this approach in a study of adjective check list response. They found substantial support for it, and in particular for a "vector" form for the function:

$$x_{ij} = \Sigma_m w_{mi} v_{jm} + a_i$$

in which w_{mi} represents the weight which an individual attaches to dimension i in determining his endorsement, and a_i is a constant representing the individual's overall tendency to endorse items. Another way to interpret the function is that there is a direction in the meaning space that corresponds to increasing acceptability of the items for the individual; the farther an item lies in that direction, the greater the degree or probability of endorsement. Individual differences in conventional inventory scores then reflect individual differences in the parameters of this function.

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 1, No. 1 Winter 1977 pp. 41-49

An interesting aspect of this approach is the interpretation it provided for social desirability responding. Cliff *et al* (1973), showed that under conditions designed to promote socially desirable responses, there was a marked tendency for there to be a large weight in the function for the dimension in the space that corresponded to desirability. Under conditions designed to promote candidness in response, this was much less often true, although some individuals still gave a large weight to the desirability dimension.

Thus there is empirical support for the intuitively plausible model which represents inventory responding as a specific mathematical function of a cognitive schema representing the semantic meaning of the items.

The purpose of the present study was to explore the extension of those findings to inventory items that are more complex than single adjectives. That is, the question investigated was the degree to which the responses of an individual to conventional inventory items are a function of the items' positions in a meaning space determined by the independent multidimensional scaling of item-similarity judgments.

Method

The overall procedure employed in studies of this type requires several types of responses from subjects and several types of analysis of them. First, for a relatively small set of items, paired judgments of similarity of meaning are obtained. These are analyzed by an MDS procedure to derive the coordinates of the items in a meaning space. There may also be some single stimulus judgments to help interpret the dimensions. Inventory responses to the items are also obtained, sometimes under more than one type of instructional set. For each individual a multiple regression analysis is performed, in which the dependent variable is his degree of endorsement of the item and the independent variables are the items' loadings on the dimensions derived from MDS. The specifics in the present case are described below.

Stimuli

The intent of the present study implies that the items used should be unambiguous in interpretation and carefully constructed and studied so as to involve a limited number of dimensions. While the approach of the present study was not factor analytic in the usual sense, items which are factorially pure and which represent only a limited number of factors would appear to be most suitable for the present uses. The Comrey Personality Scales (CPS; Comrey, 1968) appeared to be highly suitable since they have been carefully constructed and selected to have these characteristics.

The CPS includes seven scales, and the scales in turn are composed of six to ten subscales, each having two to six highly homogeneous items. The subscales are called FHIDs for "factor homogenous item dimensions." Within a FHID, there are equal numbers of positively and negatively stated items.

Nine items from each of three scales, Shyness, Neuroticism, and Empathy, were selected for use in the present study. The items are listed in Table 1. From each scale, there are six positively stated and three negatively stated items listed, the latter being as parallel as possible to the three positive items that they immediately follow in the table.

These 27 items formed the main focus of the study. They were used in two overlapping subsets of 18 each, six from each scale. In one subset (All Positive) the 18 consisted of the six positive items from each scale; in the other (Mixed) it consisted of the first three positive items and the three negatives from each scale. It was thought that multidimensional scaling of the All Positive items might yield three clusters of six items each, while the Mixed scaling would yield three bipolar dimensions. In the All Positive case, the desirability dimension was expected to be greatly truncated.

Instruments

There were two questionnaires used in the study. The first consisted of all the items from

Table 1

Items Selected from Comrey Personality Scales

Shyness (Social ease)
S1. I feel comfortable with people I have never even seen before.
S2. It is easy for me to talk with people.
S3. If I were applying for a job, I would enjoy having a personal interview.
S4. I am a talkative person.
(–) I do less than my share of talking in a conversation.
S5. At a party I like to meet as many new people as I can.
(–) I try to avoid contacts with new people.
S6. If I think I recognize someone, I will ask him if we've met before.
(–) I dislike being thrown together with people I don't know very well.

Empathy (Selflessness)
E1. I enjoy helping people even if I don't know them very well.
E2. I am a very sympathetic person.
E3. I am willing to share with others less fortunate.
E4. I have a strong desire to do something for the good of humanity.
(–) I would try to avoid a job in which I had to help people with their problems.
E5. I prefer to look after the welfare of the ones I love before I worry about myself.
(–) I take care of myself before I think about other people's needs.
E6. I find that people are fascinating.
(–) I find it difficult to get interested in what is happening to other people.

Neuroticism (Positive affect)
N1. I am successful in what I do.
N2. I feel able to deal with the problems I face.
N3. When I look back, I think that life has been good to me.
N4. I expect things to turn out for the best.
(–) Whan I want something to happen, I have the feeling that it won't.
N5. I relax without difficulty.
(–) My nerves seem to be on edge.
N6. I think I am just as good as the people I know.
(–) I have the feeling that the people I know are better than I am.

the three CPS scales (not just the nine from each listed in Table 1) plus some repeats, for a total of 115 items. The items were in random order. The response scale for these inventories was the four-point scale used by Cliff *et al.* (1973). There were two kinds of directions for the inventory: One instructed the subject to respond candidly (Candid) while the other instructed him to respond as he would if applying for a job as a management trainee (Faking).

The second questionnaire was the basis for the MDS. It included all 153 pairs of the 18 items in either the Positive or the Mixed sets described above. Subjects were instructed to judge the similarity of meaning of the members of each pair on a nine-point scale from "extremely

similar" to "extremely different." The 18 items judged for similarity were also rated twice for favorableness, once from the subject point of view and once from his estimate of a job interviewer's point of view.

Subjects

Subjects were students in an introductory psychology course fulfilling a course requirement that they participate in research.

Procedure

Subjects were tested in small groups. They completed the inventory under *both* Candid and Faking directions, using separate booklets, then *either* the Positive or the Mixed judgment booklet. There were 47 subjects in the Positive group and 44 in the Mixed. Matching of a subject's booklets was made possible while preserving response anonymity by having each subject make up his own identification number and write it on each of his booklets.

Analyses

There were three stages in the analysis. As a preliminary step, a between-persons cluster analysis (Veldman, 1967) of the similarity judgment data was performed in order to identify subgroups with potentially different perceptions of the item meanings. Next, nonmetric MDS analyses (TORSCA; Young and Torgerson, 1967) were performed on the average similarity ratings of the pairs by the members of each of the person-clusters and also on the average for all the subjects in the Positive or Mixed groups. Finally, four multiple regression analyses were performed for each subject. The dependent variables were his degrees of endorsement (1 to 4) of the items under Candid and under Faking directions. The two sets of independent variables were the projections of the items on the dimensions derived from the MDS analyses of the similarity judgments from his cluster or his total group (Positive or Mixed). These analyses would determine the degree of relation between the meaning dimensions and the inventory responses.

Results

Cluster Analyses

Results of the cluster analyses of similarity judgments suggested that there were six subgroups of persons in each of the two groups plus two or three individual mavericks. However, in the subsequent MDS and regression analyses there were almost no qualitative differences in interpretation when the dimensions for the two total groups were used instead of those for the cluster of which the individual was a member. Therefore, in the interest of brevity the rest of the results will be reported using only total group dimensions.

Nonmetric MDS

For the group judging the Mixed items, there appeared to be three dimensions, the stress value being .048. As can be seen from Figure 1, the first dimension, corresponding apparently to a desirability dimension, is larger than the second. The third, which is not shown, is slightly smaller than the second. Thus, the structure is one in which the three positive and three negative items from each scale form opposite ends of three closely related continua.

Three dimensions also seemed appropriate for the Positive group, although here the stress value was higher (.083), and the third dimension was unimportant compared to the first two. From Figure 2 it can be seen that the three clusters of six items are what accounts for the first two dimensions. The third was largely accounted for by three items, N6 and S6 at one end vs. N5 at the other.

The identity of the first dimension in the Mixed group as desirability is emphasized by the fact that the mean favorableness ratings correlated .99 with the loadings on the first dimension, in both the Candid and Faking conditions. The effect on desirability of dealing with only positive items is shown by the fact that in the Po-

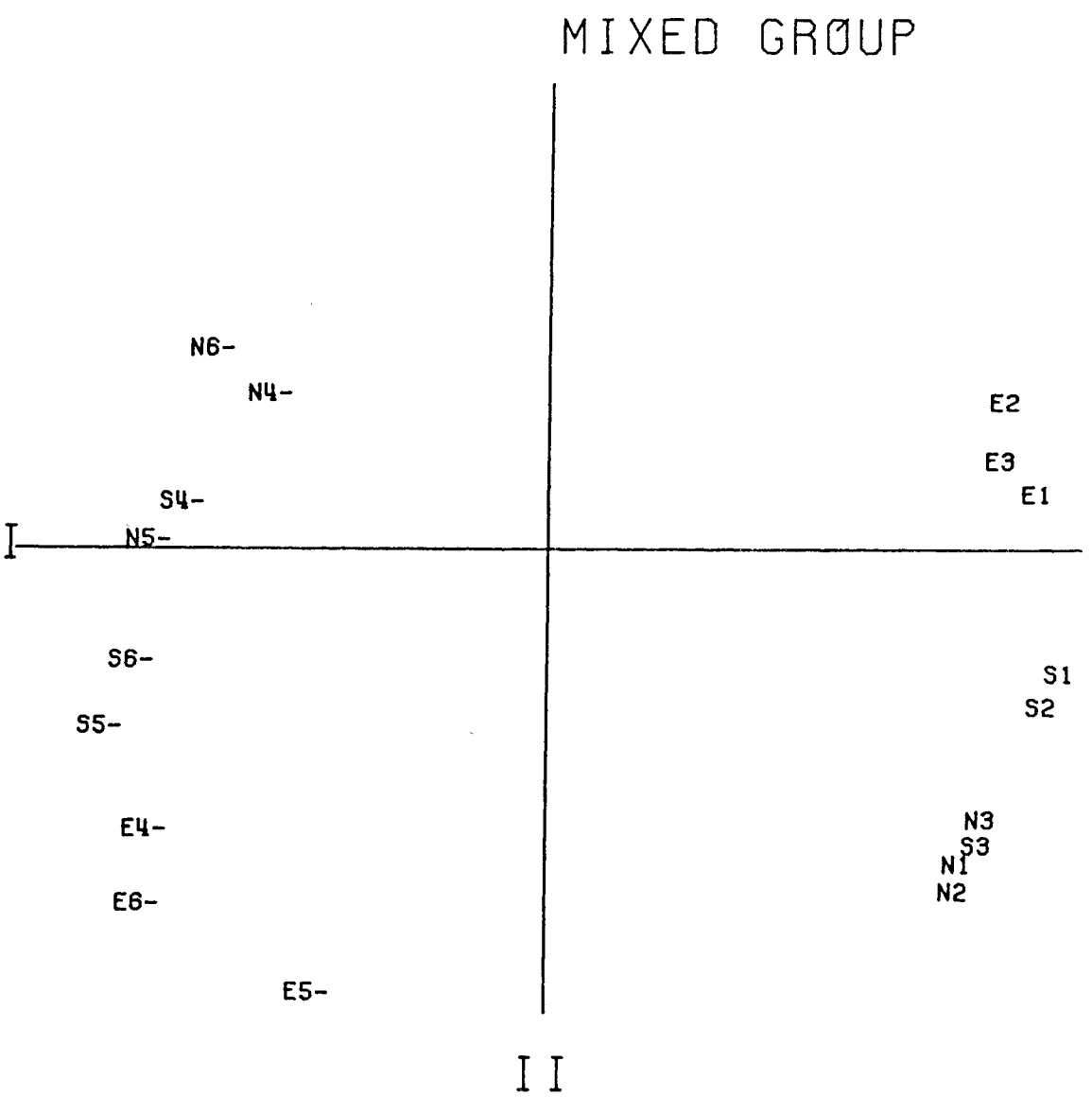

Fig. 1. First two dimensions for the item set containing both positively and negatively worded items.

sitive the first dimension correlated only – .67 ($p < .01$) with favorableness in the Candid condition and .33 (ns) with favorableness in the Faking condition. In the latter case, there was a significant correlation of – .53 with the *second* dimension. Comparing the favorableness ratings of the nine positive items in the Mixed group to the corresponding ratings in the Positive group, a surprising degree of consistency is noted, the ratings correlating .94 for the Faking condition and .62 for the Candid condition.

By including only positive items it is as if the first dimension were eliminated and we are looking at the second and third dimensions of the solution from the Mixed group. The presence of the desirability dimension in that data, though, accounts for so much of the judgment variance that the nature of the other dimensions may be somewhat obscured.

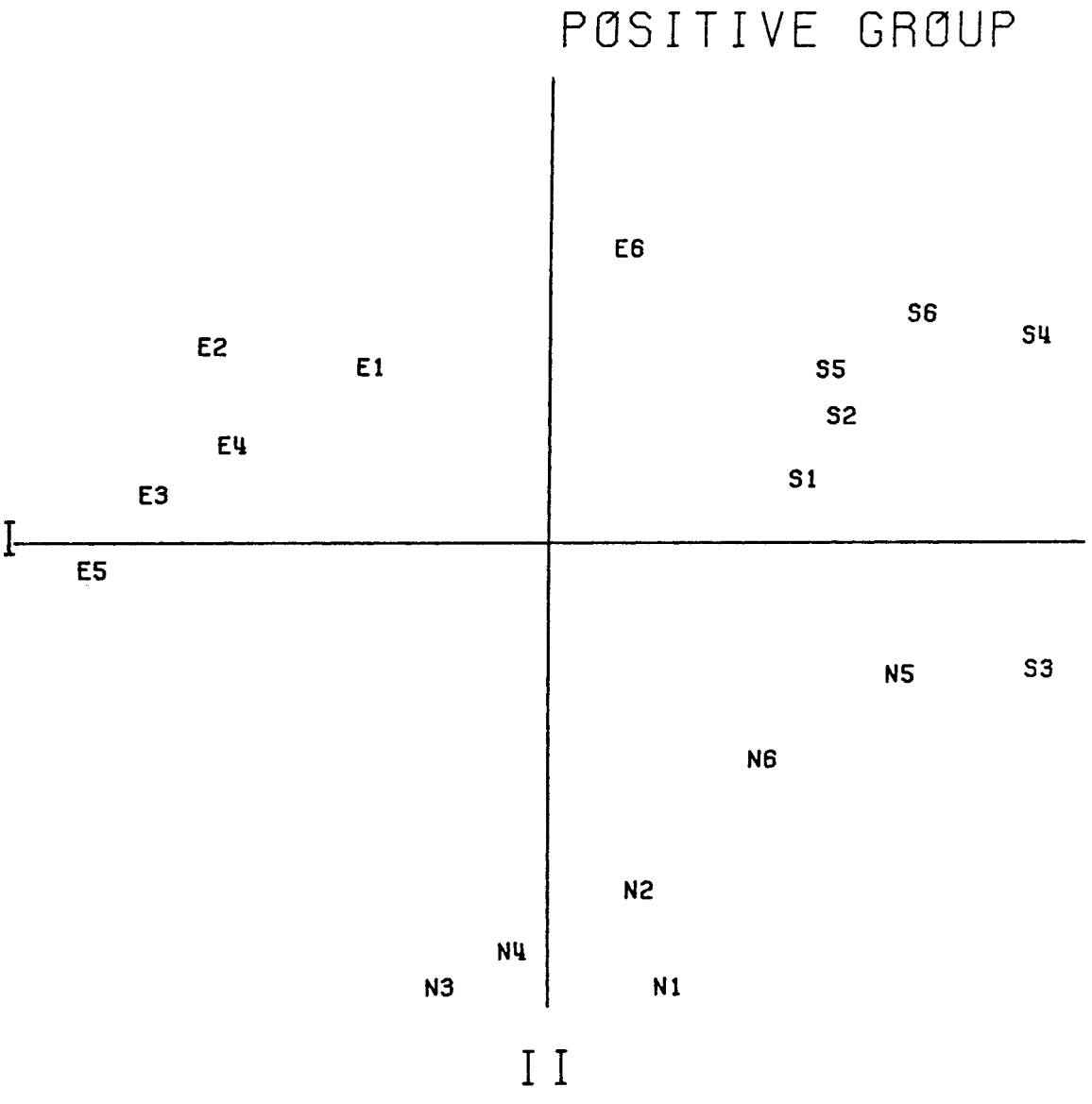

Fig. 2. First two dimensions for the item set containing only positively worded items.

Regression Analyses

The preceding results in effect only provide a background for the regression analyses which attempt to predict an individual's relative endorsement from the items' coordinates. For each individual, then, two multiple regression analyses were carried out, one for Candid endorsements and one for Faking. In these analyses, his degree of endorsement of the 18 items, expressed on a 1 to 4 scale, were the dependent variables. The independent variables in both cases were the three-dimensional coordinates of the items in the MDS space for his group (Mixed or Positive).

These regression analyses were fairly successful in the case of the Mixed group. The median unbiased (i.e., "shrunken") R for the Candid responses was .66. Of the 44 subjects in this group, 21 had R's significant at the .01 level, and eight more were significant at .05, leaving only 15 non-significant, 34 per cent.[1] This level of prediction, while somewhat lower than that reported by Cliff *et al.* (1973), is quite high when one remembers that the dependent variable is an individual's response to single items. However, in the majority of cases of significant prediction, the relation seemed attributable to the presence of a significant positive correlation with the first dimension, i.e., using desirability.

The results in the Faking condition were much stronger. There, the median unbiased multiple R was .93; ten of them were .99 or above and the lowest was .44. The latter and one of .51 were the only nonsignificant ones. Here, as expected, the influence of the desirability dimension was very strong, there being only one instance of a significant R without a significant correlation with the first dimension. Thus, when set to present a favorable image, individuals overwhelmingly responded in terms of the desirability dimension in an inventory.

The results for the Positive group were not as positive. The median unbiased multiple R for Candid responses was .36, and only 15 of 44 persons had significant multiple Rs.[2] For the Faking responses, the median unbiased multiple R was only .26, which was rather surprising. It seemed to be explained largely by the undifferentiated response to these items, many individuals endorsing almost all of these rather desirable items to the fullest possible degree. Thus there was little variance in an individual's response to predict. Apparently the substantial removal of the desirability dimension had the effect of greatly reducing the ability to relate inventory response to meaning dimensions.

Discussion

This study can be viewed as an attempt to replicate the Cliff *et al.* (1973), findings that an individual's differential response to adjective check list items can be explained largely by a functional relation with the items' coordinates in a semantic meaning space. Essentially the same methodology was followed as in the Cliff *et al.* study with the exceptions that the items were of the more common sentence type rather than single adjectives, and a group which scaled only Positive items was included. The results must be considered much less supportive of the vector model (Equation 2) in the case of Candid responding, but they strongly support the idea of the shift to using it with a high weight for the desirability dimension in the case of Faking set.

No ready explanation of the relative lack of success in predicting Candid responses seems available. Multidimensional scaling solutions were good, in the case of the Positive group as well as the Mixed. Not only was the stress low, but the clustering of items from particular CPS scales was fairly clear, indicating that the meaning of the items was fairly well described by the MDS spaces. Examination of responses to repeated inventory items indicated that an individual was quite consistent in the degree to which he endorsed a given item, so the results

[1]There were two additional cases in which there was quite strong evidence from the regression analyses for curvilinear relations.

[2]There were two additional cases in this group also in which there was quite strong evidence from the regression analyses for curvilinear relations.

are not explainable in terms of unreliability. Consistency data of the more conventional individual differences kind were also secured. It will be remembered that the individuals responded to all 30 or so items on each of the three CPS scales used, not just to the 27 items which were scaled. Considering the positively stated and negatively stated items separately, coefficient alpha (Cronbach, 1951) for the six sets varied from .71 to .85. Moreover, the correlations between total scores on the positively stated items of a scale and the negatively stated items of the same scale were .72 to .91. Thus individual differences were quite consistent.

This makes the results with Candid responses somewhat paradoxical. According to MDS, the items which are scored on a particular scale are located quite close together in a meaning space. For an individual to achieve a relatively high score on a scale, he must tend to endorse those items more than other individuals do. But within an individual, the rather weak results of the multiple regression analyses indicate that the tendency for him to endorse items in one location rather than another is usually not strong enough to show up clearly unless desirability is involved. These three results seem rather inconsistent although they are not in direct logical conflict.

The resolution of the paradox probably lies in quantitative considerations concerning the degree of the relationships involved. The substantial internal consistency reliabilities for the scales are based on a larger number of items than were used in the MDS analyses, so they would not reflect the same proportion of systematic variance. Also, "acquiesence," i.e., differential tendency to endorse items regardless of content, would contribute to the internal consistency values. Finally, it should be remembered that the regression analyses are dealing with within-person variance and the reliabilities refer to between-person variance. Thus the results are not in direct conflict.

The present study has attempted to put inventory response into the context of cognitive processing of a particular variety, one in which the response is presumed to relate in a formal way to an internalized map for the stimuli. This represents a quite different approach from that studied in such recent investigations as Roger's (in press) work, which is much more in the tradition of contemporary information processing research, and Kuncel's (1973), which is more of a psychometric approach. While the data indicate that the present model has some relevance to inventory responding, the relations are not strong enough to conclude that it forms at all a complete description of how the individual responds.

One difference between the single-adjective stimuli studied more successfully by Cliff *et al.* (1973) and the longer items of the present study is that the latter have a substantial denotative content. That is, they consist not just of the abstraction represented by a single trait adjective such as "shy," but rather of a more or less factual statement such as "I try to avoid contacts with new people." Quite conceivably there are more influences affecting an individual's endorsement of the longer statement than the degree to which it characterizes shyness or introversion as a trait, even though endorsing it might correlate somewhat with endorsing other items describing shy behavior. If nothing else, various circumstantial aspects of a person's life which affect the degree to which one item was objectively true of him might be unsystematically related to whether a second shyness item was true. That is, the longer items carry considerable semantic content over and above that represented by a trait or factor. Thus the apparent good fit in the MDS analyses is perhaps misleading. It is the result of subjects' tendencies to consider only common dimensions of meaning of judging similarity, whereas other factors come into play in inventory response.

This is much less true of single trait adjectives. Therefore, it perhaps should not be surprising that the relation between degree of endorsement and location in the MDS space was lower in the present case, at least for candid endorsements. The desirability responding seems consistent with this. Under that set, the typical individual simply abstracts a single aspect of the

item, its favorableness or desirability, and responds to this rather than to the item's content; here the relations are high. But when responding candidly, denotative as well as connotative aspects of the item influence his degree of endorsement. Thus endorsement is not as closely related to the two or three meaning dimensions uncovered by MDS as it would be in the case of the single adjective with its relatively simple meaning which is reasonably well-captured by the MDS results.

References

Cliff, N., Bradley, P., and Girard, R. The investigation of cognitive models for inventory response. *Multivariate Behavioral Research,* 1973, *8,* 407-425.

Comrey, R. L. *The Comrey personality scales.* Los Angeles; UCLA, 1968. Mimeograph.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika,* 1951, *16,* 297-334.

Kuncel, R. B. Response processes and relative location of subject and item. *Educational and Psychological Measurement,* 1973, *33,* 545-564.

Rogers, T. B. An analysis of two central stages underlying responding to personality items: the self referent decision and response selection. *Journal of Research in Personality,* In press.

Veldman, D. J. *Fortran programming for the behavioral sciences.* New York: Holt, Rinehart, and Winston, 1967.

Young, F. W. and Torgerson, W. S. TORSCA, a FORTRAN IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science,* 1967, *12,* 498.

Acknowledgements

This research was supported by PHS research grant MH-16474 from the National Institute of Mental Health. The assistance of Dr. Roger Girard in all phases of the study is gratefully acknowledged.

Author's Address

Norman Cliff, Department of Psychology, University of Southern California, Los Angeles, CA 90007.

COMPUTER PROGRAM EXCHANGE

(Continued from page 40)

EXPAK: A FORTRAN IV Program for Exploratory Data Analysis

David Thissen and Howard Wainer, The University of Chicago

Description. EXPAK, an exploratory data analysis package, is designed to facilitate exploratory data analysis by "jotting down" batches of data in various ways so that the data analyst may see things in the data which he might otherwise miss. With this goal in mind EXPAK produces the following output, parts of which may be included or suppressed, depending on their relevance to the problem at hand: (1) stem and leaf displays of each variable (a kind of labeled histogram; Tukey, 1970) with conventional and third-generation summary statistics; (2) gapping analysis of each variable, in which unusual data gaps are noted; (3) a potential transformation of the data for each variable, chosen with the goal of equalizing the tails of the distribution of that variable as measured from the median and from the stem-and-leaf and gapping displays of the transformed variates. If there is more than one variable, the program also prints the correlation matrix of the untransformed variables and two-way scatter plots, as well as plots of the variables transformed with the goal of "straightening" the relationship between one variable and another. Individual data points are identi-

(Continued on page 50)