Implicit Theories of Personality: Further Evidence of Extreme Response Style

Willem A. Van der Kloot, Pieter M. Kroonenberg, and Dini Bakker
University of Leiden

Five groups of 19 subjects made ratings on 11 personality trait scales of overlapping subsets of 59 artificial stimulus persons who were described by one to five personality trait adjectives. The data were analyzed per group of subjects (blockwise) and per type of stimulus person (questionnaire-wise) by means of three-mode principal component analyses. This yielded highly similar structures for the scales, and in the blockwise analyses, for the stimulus persons. This similarity was substantiated by external three-mode analyses, which showed that all stimulus persons can be mapped into one overall configuration. In all analyses it was found that differences between subjects consisted of differences in extremity of judgment, which suggests the operation of response style.

In a recent paper on individual differences in implicit theories of personality, Van der Kloot and Kroonenberg (1982) presented evidence that the perceived relations between 11 personality traits can be described adequately by mapping them on a circular contour. Moreover, they found that individual differences among the subjects were restricted to one dimension, which means that the subjects differed from each other not with regard to the shape of their trait structures, but mainly with regard to the size of these configurations. These results suggest the operation of individual response styles (i.e., with regard to extremity of judgment) instead of idiosyncratic cognitive structures.

The finding that personality traits can be summarized by means of a circumplex fits in nicely with the results of other studies. Circular representations of personality traits were also obtained by Becker and Krug (1964), Benjamin (1974), Conte and Plutchik (1981), Lorr and McNair (1963), Lyons, Hirschberg, and Wilkinson (1980), Rinn (1965), Sherman (1972), Tucker (1972), Van der Kloot and Slooff (in press), and Wiggins (1979) (see also Wiggins, Steiger, & Gaelick, 1981). The structures found in those studies are described in terms of a love-hate or evaluation dimension and a dominance-submission dimension and thus are quite similar to the circumplex (or rather radex) model proposed by Leary and his colleagues (Freedman, Leary, Ossorio, & Coffey, 1951; Leary, 1957) and by Foa (1961).

The data presented by Van der Kloot and Kroonenberg (1982)
were obtained in an impression formation experiment in which subjects had to rate stimulus persons who were described by one or two personality trait adjectives. The present paper reports an attempt to generalize the findings of Van der Kloot and Kroonenberg to different types of judgments, that is of stimulus persons described by more than two adjectives.

In particular, we wanted to study (a) the shape of the trait scale space, (b) the structure of possible individual differences, and (c) the configurations of the stimulus persons. In view of our earlier results we expected to replicate the circular shape of the trait scale space, and the uni-dimensionality of the subject space. The individual difference component was expected to again indicate the extremity of the subjects' response styles. Finally, we wanted to obtain one common representation of all stimulus persons, so that the relations between the one-, two-, three-, four-, and five-adjective stimulus persons can be studied.

Method

Stimulus Material

The stimulus material consisted of eleven personality trait adjectives: likeable, cooperative, intelligent, industrious, dominant, aggressive, unreliable, pessimistic, passive, submissive, and modest. Earlier research (Van der Kloot & Kroonenberg, 1982; Van der Kloot & Van den Boogaard, 1978) has shown that these adjectives lie on a circle in the order in which they are mentioned above.

These stimuli were used to construct five questionnaires that differed with regard to the type (and number) of stimulus persons described. Questionnaire 1 contained the descriptions of 11 stimulus persons. Each stimulus person was described by one of the adjectives (e.g., somebody is aggressive). Questionnaires 2, 3, 4, and 5 each contained 12 stimulus persons who were described by combinations of two, three, four, and five adjectives, respectively. Each description was printed on top of a different page and was followed by eleven 10-point rating scales. The rating scales were labeled with the 11 personality traits mentioned above, including the adjective(s) used in the description of the stimulus person. The rating scales ranged from 1 to 10, with end points denoted by "extremely not . . ." and "extremely . . ." (e.g., "extremely not cooperative" and "extremely cooperative"). The combinations of two, three, four, and five adjectives that were used in these
questionnaires (see Table 1) were selected on the basis of results by Van der Kloot and Van den Boogaard (1978). These authors found a

Table 1

*Stimulus Persons Described by Combinations of Adjectives*

<table>
<thead>
<tr>
<th>(Questionnaire 2)</th>
<th>(Questionnaire 3)</th>
<th>(Questionnaire 4)</th>
<th>(Questionnaire 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-adjective Stimulus Persons</td>
<td>Three-adjective Stimulus Persons</td>
<td>Four-adjective Stimulus Persons</td>
<td>Five-adjective Stimulus Persons</td>
</tr>
</tbody>
</table>
Figure 1.
The configurations of the scales in the five blockwise analyses
circular trait configuration almost identical to the one presented in Figure 1. If the traits are represented as vectors from the origin, the relationships between pairs of traits can be described by the angles $\phi$ between the vectors. These angles were divided in five classes: $0^\circ$ with $0^\circ \leq \phi < 22^\circ$, $45^\circ$ with $23^\circ \leq \phi < 67^\circ$, $90^\circ$ with $68^\circ \leq \phi < 112^\circ$, $135^\circ$ with $113^\circ \leq \phi < 157^\circ$ and $180^\circ$ with $158^\circ \leq \phi \leq 180^\circ$. Combinations of traits to describe stimulus persons were selected such that all classes were represented. For instance, the five-adjective combination likeable-dominant-intelligent-industrious-passive contains one $0^\circ$, four $45^\circ$, two $135^\circ$ and two $180^\circ$ relationships.

Because we believed that rating 59 stimulus persons on 11 rating scales would probably be too tiring and boring for the subjects and thus would lead to unreliable data, we made five different tasks, each task consisting of a different combination of three questionnaires. The combinations of questionnaires used were 1-2-3, 1-2-4, 1-2-5, 1-3-4 and 1-3-5. The questionnaires within each task were presented in two booklets each preceded by an instruction page. One booklet contained the first questionnaire, the second booklet contained the combination of the remaining questionnaires. Four different random orders of the 11 rating scales on a page were used and of each type of booklet four different forms were constructed by varying the order of the pages within the questionnaires in a random manner.

Subjects

The five types of questionnaires were divided evenly (and at random) over the 95 subjects who participated in this study. The subjects were recruited through advertisements in the weekly Leiden University newspaper, and by means of posters in a number of university buildings. With 100 persons who volunteered to participate, individual appointments were made for one of four possible sessions. Of these 100 persons 5 failed to show up.

Procedure

The subjects arrived at individually convenient times in a large classroom where they received the booklets corresponding to one of the five tasks. After completion of the questionnaires, which took between 60 and 90 minutes, the subjects received 10 Dutch guilders (approx. $4.00).
Table 2

Allocation of Questionnaires to Blocks of Subjects

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1 (n = 19)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (n = 19)</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>3 (n = 19)</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>4 (n = 19)</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>5 (n = 19)</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

*a* The questionnaire numbers correspond to the number of adjectives used for describing the stimulus persons

**Data**

The data collected in this study form an unbalanced incomplete design of the form displayed in Table 2. They can be regarded *blockwise* as five three-way/three-mode blocks of 35 stimulus persons (Mode 1) by 11 scales (Mode 2) by 19 subjects (Mode 3), or *questionnaire-wise* as five three-way/three-mode matrices of respectively 11, 12, 12, 12, and 12 stimulus persons by 11 scales by 95, 57, 57, 38, and 38 subjects. Missing values, of which there were 36 (on a total of 36575 ratings), were replaced by the mean ratings (i.e., averaged over the subjects within a block) of the particular stimulus person on the particular scale. As a preliminary step in the analysis, the ratings

1Although there are so few missing data that even replacing them by random numbers would not have a noticeable effect, we chose the present manner of substitution because it preserved the (mean) interactions of the particular stimulus-scale combinations. Any other number could introduce unwanted individual differences.
were double-centered per subject in order to remove some unwanted sources of variation, notably the "main effects" per subject of scales and stimulus persons. These effects were removed because they constitute additive components that we did not want to include in the multiplicative model that is fitted by TUCKALS3 (see, for a discussion of various possibilities for centering three-way data, Harshman & Lundy, 1984; Kroonenberg, 1983, p. 127; Kruskal, 1981).

Analyses

The data were analyzed by means of TUCKALS3, an alternating least squares program for three-way principal component analysis. This program was developed by Kroonenberg and De Leeuw (1980) and Kroonenberg (1983). For a stimuli by scales by subjects data-matrix TUCKALS3 derives components for the stimuli, scales, and subjects, and computes a core matrix that describes how the components are related. In contrast to the original method of Tucker (1966), the Kroonenberg and De Leeuw method enables a partitioning of the total sum of squares into a fitted sum of squares, SS(FIT), and a residual sum of squares, SS(RES).

The model that is fitted can be written as

\[ \hat{Z}_k = G \left( \sum_u e_{ku} C_u \right) H', \]

where \( \hat{Z}_k \) is an approximation of the stimuli by scales matrix of subject \( k \). \( G \) and \( H \) are component matrices for the stimuli and scales, respectively, \( C_u \) is the core matrix corresponding to the \( u \)-th subject component, and \( e_{ku} \) is the loading of subject \( k \) on this component. If there is only one substantial subject component, this formula reduces to

\[ \hat{Z}_k = e_k GCH', \]

which shows that the subject loadings, \( e_k \), act as a proportionality factor by which the common stimulus and scale configurations are enlarged or shrunk.

Recently TUCKALS3 has been extended with a fix option (Kroonenberg, Van der Koot, & Brouwer, 1983; Van der Koot & Kroonenberg, in press) which enables one to fix one or two of the stimulus, scale, and subject configurations, and derive component loadings for the remaining way(s) of the data.
Results

The centered data were initially analyzed per block of subjects. Such a blockwise analysis has the advantage that one analysis includes three types of stimulus persons and thus may yield a more general and stable scale space in which 35 stimulus persons can be mapped. It also has the advantage that each subject occurs in only one analysis. It has the disadvantage that different analyses may yield similar results because of the overlap among the stimulus persons, in particular, because in all analyses the same 11 one-adjective stimulus persons are included. In order to check whether such "artifacts" occurred, the data were also analyzed questionnaire-wise.

Blockwise Analyses

For each $35 \times 11 \times 19$ block a solution consisting of two components per mode was derived. The fit of a solution to the data can be expressed as a proportion explained variation, i.e., the ratio of the fitted sum of squares ($SS(FIT)$) to the total sum of squares ($SS(TOTAL)$). These fit measures ranged from .50 for the fifth block to .57 for block one and thus are comparable to the .51 obtained in the $31 \times 11 \times 59$ problem described by Van der Kloot and Kroonenberg (1982). As Kroonenberg has shown (1983; see also Van der Kloot & Kroonenberg, 1982) the total fit can be partitioned into separate contributions of each component per mode. These contributions and the total fit are listed in the upper part of Table 3.

From this table it is clear that the total fitted sums of squares of all blocks are comparable and that corresponding components in different blocks explain approximately equal amounts of variation. It is also clear that the explained variation of the first component in Mode 3 is much larger than that of the second component. The ratio of the contributions of the two components in Mode 3 varies between $82:1$ and $102:1$, suggesting that the subject mode in each block is essentially one-dimensional. These results are almost identical to those of Van der Kloot and Kroonenberg (1982).

The results for the different blocks are not only similar with respect to the third mode, but also lead to similar interpretations of the components in Mode 1 and 2. In all blocks one can clearly identify an evaluation dimension and a dominance dimension. These dimensions and the roughly circular pattern of the scales in these spaces (see Figure 1) are remarkably similar, both to each other and to the results of Van der Kloot and Kroonenberg (1982). However, as was stated
Table 3

Proportion of $SS(FIT)$ to $SS(TOTAL)$ of two Components in three Modes for Blockwise, Questionnaire-wise and External Analyses

<table>
<thead>
<tr>
<th>Mode</th>
<th>Blockwise analyses</th>
<th>Questionnaire-wise analyses</th>
<th>External analyses with scales fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>.35</td>
<td>.22</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>.35</td>
<td>.21</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>.31</td>
<td>.21</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>.33</td>
<td>.21</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>.28</td>
<td>.22</td>
<td>.28</td>
</tr>
</tbody>
</table>

Van der Kloot & Kroonenberg, 1982

.26 .25 .26 .25 .498 .015 .51
above, this similarity may be the result of the considerable overlap among the stimulus persons in the different blocks, in particular of the 11 one-adjective stimulus persons. To check this possibility the data were also analyzed questionnaire-wise, which has the advantage that a solution for one type of stimulus persons is independent of the solutions for the other types. It has the disadvantage that the same subject occurs in three different analyses, which introduces another kind of dependency between the results.

**Questionnaire-Wise Analyses**

For the five three-mode matrices of dimensions $11 \times 11 \times 95$ (Questionnaire 1), $12 \times 11 \times 57$ (Questionnaire 2 and 3) and $12 \times 11 \times 38$ (Questionnaire 4 and 5) solutions with two components per mode were obtained. The fitted sums of squares, which are displayed in the middle part of Table 3, range from .45 (Questionnaire 5) to .62 (Questionnaire 2). In all analyses the fitted sums of squares of the corresponding components of the stimulus and scale spaces are almost equal and the corresponding components of the subject spaces have similar contributions. The ratio of these contributions varies from 438:9 to 611:7, which again shows the one-dimensionality of the subject mode. Again the circular shape of the scale space is clearly recognizable (see Figure 2), although there are differences among the five analyses with regard to the particular pattern of the individual scales along the "circles."

One way to test whether the scale spaces of the five analyses can be represented by one common configuration is to subject the data to an external or constrained analysis in which the scale space is fixed to an a priori configuration. We used TUCKALS3 again to analyze Questionnaires 2 through 4 with the scale configuration fixed to that of Questionnaire 1. This scale space was chosen because one could regard this as the most "pure" scale space, in the sense that it is uncontaminated by the specific combinations of adjectives that were selected. Moreover, this scale space was believed to be more stable as it was derived from the ratings of 95 (instead of 57 or 38) subjects.

The fitted sums of squares of these external analyses are represented in the lower part of Table 3. Going from Questionnaire 2 to Questionnaire 5 there is a decrease in fit, as was to be expected. As this decrease is never more than .07, the conclusion seems warranted that the fixed trait scale space is a reasonable approximation of the trait spaces found in the separate unconstrained analyses.

---

^This way of analyzing the data was suggested to us by an anonymous reviewer.
Figure 2.
The configurations of the scales in the five questionnaire-wise analyses

OCTOBER 1985
Both in the blockwise and the questionnaire-wise analyses we found subject spaces that are essentially one-dimensional. As Van der Kloot and Kroonenberg (1982; in press) have shown, one-dimensionality of the subject space means that—within the model—the shape of the individual stimulus and scale configurations are the same for all subjects, and that these configurations only differ with regard to size. In such cases, each subject can be represented by one parameter, her or his loading on the first component of the subject space, which is a measure of the size of the configuration of stimulus persons and scales as used by the subjects.

In the 1982 study of Van der Kloot and Kroonenberg, the subject loadings covaried with the sums of squares of their ratings. The same relationship was also found to exist in the present study. In the blockwise analyses the product-moment correlations between the subjects' total sums of squares and their loadings on the first component of Mode 3 were .907, .912, .955, and .951 for Block 1 through 5 respectively. In the analyses per questionnaire these correlations were .896, .903, .928, .941, and .926 for the Questionnaires 1 to 5. Therefore it may be concluded that the individual differences emerging in this study again indicate the operation of extreme response style. Subjects who tend to use the extremes of the scale more often, have larger sums of squares and have higher loadings on the subject dimension.

As suggested by an anonymous reviewer, another way to check whether extreme response style is operating, is to standardize the data matrices of all subjects so that their sums of squares become equal. Individual differences in extremity are then eliminated, and the loadings on the first subject component should be approximately constant. Questionnaire-wise analyses of the standardized data showed that there was little change in fit, but that the variance of the loadings on the first subject dimension decreased by approximately 75%. The effect of standardization on the analysis of the $11 \times 11 \times 95$ data of Questionnaire 1 is shown in the stem-and-leaf diagram of Figure 3. This supports our conclusion that the original subject component reflects extreme response style.

Joint Representation of Stimulus Persons

One objective of the present study was to obtain a common representation of all 59 stimulus persons. In order to do so, we used the results of the constrained analyses and computed for each type, $q$, of
stimulus persons \((q = 1, \ldots, 5)\) coordinates

\[
Y^{(q)} = G^{(q)} C^{(q)}
\]

which map the stimulus persons into the common scale space \(H\) of the one-adjective stimulus persons. The two-, three-, four- and five-adjective stimulus persons are represented in Figures 4 to 7. In each figure also the one-adjective stimulus persons of Questionnaire 1 (represented by the vectors) are shown. Although these figures are presented separately for reasons of clarity, they may be superimposed to compare the positions of the different types of stimulus persons. This enables one to study the implicit theories of personality for stimulus persons described by different numbers of adjectives. Preliminary results of such studies were reported by Van der KIooot (1984).

**Discussion**

The fact that Van der KIooot and Kroonenberg (1982) obtained results that were also found in other studies yielded some confidence in...
Figure 4.
Joint representation of one-adjective stimulus persons (denoted by the vectors) and two-adjective stimulus persons (denoted by the figures, which refer to Table 1)

Figure 5.
Joint representation of one-adjective stimulus persons (denoted by the vectors) and three-adjective stimulus persons (denoted by the figures, which refer to Table 1)
the generalizability of these findings. This confidence is reinforced by the present results, because they replicate those of 1982: first of all the circularity or circumplexity of the personality trait space, and secondly the unidimensionality of the subject space.

A new result is the high degree of stability of the separate solutions for the different (sub)samples of subjects who had to perform rating tasks that were partially different. It is noteworthy that in the blockwise analyses there is only a slight decrease in the $SS(\text{FIT})/SS(\text{TOTAL})$ ratio for the groups which had to rate four- or five-adjective stimulus persons, although these tasks are presumably much more complex than rating one- or two-adjective stimulus persons. This conclusion is supported by the results of the questionnaire-wise analyses. A detailed study of the structural relationships among the mappings of stimulus persons described by varying numbers of adjectives will be reported in the future.

As in the previous study there emerged only one individual difference factor, which must be interpreted as a response style factor: subjects differ with regard to the extremity of their judgments. This suggests that individual implicit theories of personality (ITPs) differ
with regard to their size rather than their form. The question remains how this result compares to the results of other studies in which individual differences were found. Of course, it is possible to find individual differences in our data by performing separate analyses on the two-mode data of each judge. However, such individual differences cannot be grouped into a relatively small number of types of judges or judgments. It seems that we either have to content ourselves with one individual difference component, or have to accept as many individual ITPs as there are subjects. On the other hand, Van der Kloot and Kroonenberg (1982) showed that even such idiosyncratic structures can be regarded as variations on a common theme, in that the "circular" shape of the scales can still be recognized if we allow for (a) more angular patterns, (b) more pronounced clusterings of scales, (c) interchanges among adjacent scales, and/or (d) shifts of separate scales to regions other than expected. Therefore, we still tend to conclude that our subjects used one common, possibly culturally shared, ITP when making their judgments.

When comparing our study with other research, it should be noted that different studies differ with regard to subjects (e.g., heterogeneous vs. homogeneous samples), stimuli (e.g., real persons vs. hypothetical, verbally described, stimulus persons), tasks (e.g., free responses vs.
fixed rating scales) and level of analysis (e.g., individual vs. multi-mode). One could speculate that individual analyses of free responses to real persons (as in Rosenberg, 1977) would show larger individual differences than three-mode analyses of fixed responses to hypothetical stimulus persons. That this speculation may indeed be true, is suggested by the results of two studies in which Tucker and Messick’s (1963) point-of-view analysis was applied to similarity ratings of personality trait adjectives (Walters & Jackson, 1966) and of real role-persons (Messick & Kogan, 1966). In the first study two subject factors were obtained of which the second described only very few individuals, whereas the second study obtained six viewpoint dimensions for females and seven points-of-view for males#. A study by Wiggins and Blackburn (1976), in which three-mode factor analysis was used to analyze the ratings of real stimulus persons by two heterogeneous groups of subjects, found two subject factors representing the members of the two groups (college fraternities). These factors can be interpreted as response style factors. The first represented raters who gave high overall mean ratings, which means that they tend to describe all stimulus persons favorably and leniently. The second dimension, on the other hand, represented raters who had “... a propensity to assign trait ratings with a high variability (i.e., raters who had a higher standard deviation [italics added] across all traits)” (Wiggins & Blackburn, 1976, p. 276). Finally, it should be noted that even Rosenberg’s (1977) results, which are based on separate analyses of free responses to real stimulus persons, show large similarities between the various individual structures, in that a general evaluative dimension appeared as a strong dimension in the structure of each subject. Moreover, Rosenberg (1977, p. 225) tentatively discerned two important areas in which traits can be categorized: acceptance and competence. In our opinion, these categories correspond to the two dimensions found in our study: evaluation and dominance. Therefore, we think it is reasonable to conclude that people use a common, underlying, frame of reference consisting of evaluation and dominance when judging other persons’ personalities. People may differ, however, with regard to the “indicants” (Wiggins & Fishbein, 1969, p. 184) of this framework; although the framework itself remains clearly recognizable.

The operation of extreme response style in person perception has been found not only by Wiggins and Blackburn (1976), but also in other studies, for example Walters and Jackson (1966) and Messick

#To us it is a matter of debate whether these large numbers of subject dimensions, and, for instance, the five subject types of Sherman (1972), are necessary or reflect attempts to over-fit the data.
and Kogan (1966). Schneider (1973) reviewed those and a number of other studies and concluded that "... there has been limited success relating such (individual) differences to traditional personality variables" (p. 305). A similar conclusion was reached within the field of personal construct psychology. In particular, the hypothesis (Bonarius, 1971; Hamilton, 1969) that the extremity of a person's ratings is related to his or her maladjustment (e.g., rigidity, anxiety and immature cognitive development) was not substantiated in a series of experiments by Bonarius (1971). He reached the conclusion that extreme responses are the result of the personal relevance of the stimuli and scales (constructs) that are used, and of their interaction. This theory could explain the present findings, if we assume that there were large differences among our subjects with regard to the personal relevance of our stimulus persons and our rating scale adjectives. However, we have no means to substantiate this hypothesis, as the pertinent questions were not incorporated in our questionnaires. Adding such questions in future research would seem to be worthwhile.

References

Willem A. Van der Kloot et al.


