What is Cognitive Complexity? Alternative Scoring Procedures for the Modified Rep Test

Tom Trabin, Abbott Northwestern Hospitals
Phillip Wood, Search Institute

The way people think about and thereby act upon their environment is partly dependent upon how they differentiate between the various aspects of that environment. Individual differences along these lines help to explain how some people may overlook critical elements in dealing with a given situation. For example, a supervisor must be able to distinguish between the several relative strengths and weaknesses of a job candidate in coming to a decision about employment. While a candidate's personal habits and likeableness may be kept in mind in making a decision to hire, these aspects should be kept distinct from other strengths and weaknesses in order to come to a well-reasoned decision. Other ready examples come from education. In student evaluation of instructors, it is vital to understand the degree to which student perceptions of some instructor traits, such as amiability, influence other ratings of the course, such as the fairness of grading and adequacy of course material. Clearly, assessment of an individual's construing of the environment is important for understanding the actions of that individual in a given setting.

In this paper we shall discuss various approaches to scoring an instrument which purports to measure the degree to which an individual is able to perceive a given environment complexly. Two new approaches to scoring the instrument will be contrasted with the current scoring method. Individuals' scores under each of the three methods will then be correlated with scores from three other measures of complexity.

Background

The degree to which people differentiate between various aspects of their environment has been studied with respect to such diverse stimuli as paintings, types of bread, and various inanimate objects (Bannister & Mair, 1968) as well as cigarettes (Brotzge & Crain, 1977) and flowers and automobiles (Cohen & Feldman, 1982). Within the field of social psychology, the degree to which persons differentiate between diverse components of their social environment has been termed cognitive complexity. Bieri et al. (1966, p. 185) has defined cognitive complexity as an information processing variable which helps to predict how an individual transforms specified behavioral information into social or clinical judgments. More specifically, cognitive complexity has been defined as the capacity to construe social behavior in a multidimensional way. In summary, most studies have conceptualized cognitive complexity as the degree of dimensionality present in an individual's judgment of stimuli.

Method

Subjects for this study were 497 male and female lower division college and university students from 18 classes in the physical sciences, social sciences, and humanities. Several measures were administered during three different time periods during an academic quarter.

The Role Construct Repertory Test (Rep Test), as modified by Bieri et al. (1966) was administered as a measure of cognitive complexity. In this test, the columns of the instrument are ten different roles (i.e., self, parent, friend, etc.) and the subject is asked to rate those roles on ten constructs (trait descriptors) using a six-point Likert scale to indicate the degree to which the given construct is true of that role. This test is frequently scored according to Bieri's algorithm (to be discussed below) in studies designed to assess degree of complexity of social perception. Three other measures administered and of relevance to this study were the Group Embedded Figures Test (Oltman, et al., 1971), the Paragraph Completion Test (Hunt, et al., 1978) and a student evaluation of instruction questionnaire developed from studies by Doyle and Whitely (1974). These measures will be explained in a discussion of the correlation of these tests with various approaches to scoring the Rep Test.

Scoring of the Rep Test

Three scoring algorithms were used to measure cognitive complexity with the Rep Test. The first, Bieri's d, is the commonly used approach. The second approach consisted of the extraction of the first characteristic root of a subject's responses based on a factor analysis of the rows of the Rep Test. The third approach consisted of the average standard deviation of ratings within the columns of the test across all roles.

Bieri's d

For a given column (i.e., role) of the Rep Test, a dissimilarity measure is computed by summing all possible pairwise comparisons of ratings for a given column with a score of 1 given for each comparison in which the ratings are different. This dissimilarity is then summed across all roles to arrive at a d score for an individual. As mentioned above, cognitive complexity is defined as the multidimensionality of an individual's perceptions. As we will now show, however, Bieri's d is not a measure of the multidimensionality of an individual's Likert ratings of roles. The sufficient statistic for the distribution of d depends, in fact, upon the conditional frequencies of Likert values within each of the columns of the instrument.
Table 1

<table>
<thead>
<tr>
<th>Scoring Algorithms</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bieri’s d²</td>
<td>55.323</td>
<td>-1.006</td>
<td>1.464</td>
</tr>
<tr>
<td>standard deviation</td>
<td>63.144</td>
<td>0.330</td>
<td>0.318</td>
</tr>
<tr>
<td>first eigenvalue</td>
<td>71.987</td>
<td>0.877</td>
<td>1.475</td>
</tr>
</tbody>
</table>

Table 2

Selected Rep Test Examples

<table>
<thead>
<tr>
<th>SUBJECT A</th>
<th>SUBJECT B</th>
</tr>
</thead>
<tbody>
<tr>
<td>roles</td>
<td>roles</td>
</tr>
<tr>
<td>c 1 2 2 2 2 2 2 2 2</td>
<td>c 1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>o 2 4 3 3 2 3 4 4 4 4</td>
<td>o 2 4 2 2 2 2 2 2 2 2</td>
</tr>
<tr>
<td>n 3 2 2 2 2 2 2 2 2 2</td>
<td>n 3 1 5 4 6 1 2 2 1 1</td>
</tr>
<tr>
<td>s 4 5 4 2 3 3 3 3 2 2 3</td>
<td>s 4 1 6 5 5 5 3 6 5 4</td>
</tr>
<tr>
<td>t 5 2 2 2 2 2 2 2 2 2 2</td>
<td>t 5 4 1 2 4 3 1 2 1 4</td>
</tr>
<tr>
<td>r 6 2 2 2 2 2 2 2 2 2 2</td>
<td>r 6 4 2 2 4 3 3 3 2 3</td>
</tr>
<tr>
<td>u 7 3 2 2 2 2 3 2 3 3 2</td>
<td>u 7 1 4 6 6 2 1 5 3 2 1</td>
</tr>
<tr>
<td>c 8 4 3 4 3 4 4 4 4 3 4</td>
<td>c 8 1 5 4 6 1 2 2 1 1 1</td>
</tr>
<tr>
<td>t 9 3 2 2 3 4 2 3 3 4 3</td>
<td>t 9 4 3 2 2 3 2 2 1 3</td>
</tr>
<tr>
<td>s 1 0 3 5 4 4 4 5 4 5 4</td>
<td>s 1 0 5 4 3 2 4 3 3 4 3</td>
</tr>
</tbody>
</table>

Figure 1

Scatterplot of Bieri's d² score by First Characteristic Root

PLOT OF O_SCORE*£IGENl LEGEND: A = OBS, B = 2 OBS, C = OBS, D = OBS, E = OBS, F = OBS, G = OBS, H = OBS, I = OBS, J = OBS, K = OBS, L = OBS, M = OBS, N = OBS, O = OBS, P = OBS, Q = OBS, R = OBS, S = OBS, T = OBS, U = OBS, V = OBS, W = OBS, X = OBS, Y = OBS, Z = OBS.
To see this, consider the case of the $i$th column of ratings for a three-point Likert scale. Let $A_1$, $A_2$, and $A_3$ represent the frequencies of Likert values 1, 2, and 3 in a given column.

Since the task of counting the dissimilarities in a given column for the 10 rows can be done without respect to order, the formula for $d$ for the $i$th row can be given as:

$$d_i = A_{1i}(10-A_{1i}) + A_{2i}(10-A_{1i}-A_{2i}) + A_{3i}(10-A_{1i}-A_{2i}-A_{3i})$$

The last right hand term, dealing with $A_3$ drops out because its multiplier is zero.

Multiplying out this equation and collecting terms yields:

$$d_i = A_{1i}(A_{1i} + A_{2i} + A_{3i}) + 10A_{1i} + A_{2i} + A_{1i}A_{2i}$$

which, on completing the square, and rearranging terms, yields:

$$d_i = (10-A_{1i}-A_{2i})(A_{1i} + A_{2i}) + A_{1i}A_{2i}$$

Since $10-A_{1i}-A_{2i} = A_{3i}$,

$$d_i = A_{3i}(A_{1i} + A_{2i}) + A_{1i}A_{2i}$$

which is $d_i = j(A_{i1} + A_{i2} + A_{i3})$

To maximize $d_i$ (using equation 3) we take the partial derivative with respect to $A_{1i}$ and $A_{2i}$, set the derivatives to zero and solve the set of simultaneous equations.

$$\frac{\partial d_i}{\partial A_{1i}} = 10 - 2A_{1i} - A_{2i} = 0$$

$$\frac{\partial d_i}{\partial A_{2i}} = 10 - 2A_{1i} - A_{2i} = 0$$

Substituting for $A_{1i}$ yields

$$A_{2i} = 10 - 2(10 - 2A_{2i})$$

Resubstitution yields $10/3A_{2i}$. Since $d_i = \frac{10}{3} A_{2i}$ equal numbers of Likert values across each column will maximize $d_i$.

Based on the above, Bieri's $d$ can be quickly computed by the following formula:

$$d = \frac{10}{3} \sum_{i,j,k} A_{ij}A_{ik}$$

where $A_{ij}$ is the frequency of the $j$th Likert value for the $i$th column, $A_{ik}$ is the frequency of the $k$th Likert value for the $i$th column, and $j,k$.

Since Bieri's $d$ is strictly a function of the frequencies of Likert values within the columns of the instrument, it cannot be considered a measure of the multidimensionality of a subject's ratings. In addition, it does not assess the degree of dispersion between ratings. In response to this, the following two scoring algorithms were advanced to measure the degree of dispersion within columns, and the dimensionality of subject's ratings.

Standard Deviation of Column Ratings

The magnitude of average dispersion between ratings within columns was computed by taking the standard deviation of the rows within a column, and averaging across all columns.

First characteristic root of trait ratings (Rows)

A principal axis factor analysis with intercorrelations was conducted separately on each subject's Rep Test. Traits (rows) were treated as variables and roles (columns) were treated as observations within the PROC FACTOR analysis. The first eigenvalue from the analysis was taken as a measure of cognitive simplicity. The greater the first eigenvalue, the more unidimensional the subject's social judgment was considered to be. This measure of cognitive simplicity (and cognitive complexity as its converse) fits quite well with Bieri et al.'s conceptual definition of cognitive complexity as the capacity to construe social behavior in a multidimensional way.

Interrelationships between Scoring Methods

Subject scores derived from each of the three scoring approaches were first compared with respect to range, skewness, and kurtosis. Scores were standardized to a mean of 50 and standard deviation of 10 to facilitate these comparisons. Results (see Table I) indicate that between the two types of dispersion scores, the standard deviation approach yielded a somewhat wider range of scores that were less skewed.

Unstandardized scores derived from each of the four algorithms were then correlated. The correlation between Bieri's $d$ and the first characteristic root was .175 ($p<.005$), correlation between Bieri's $d$ and the average standard deviation was .574 ($p<.0001$). The strong correlation between Bieri's $d$ and the average standard deviation is not surprising, given their shared characteristic of being essentially dispersion measures. However, if viewed as alternate forms to be correlated as a measure of reliability, the correlation between the two approaches is rather moderate and may indicate some important differences between them. These differences may be understood along lines mentioned previously (i.e., Bieri's $d$ is a measure of frequency of different ratings while the standard deviation is a measure of magnitude of dispersion).

The extremely low and positive correlation between Bieri's $d$ and the first root suggest that these two measures assess quite different aspects of the responses present in the Rep Test. The fact that this correlation is positive is not expected, given the fact that the first eigenvalue is a measure of cognitive simplicity while $d$ is a measure of cognitive complexity.
To better understand the reason for the obtained positive correlation, a scatter plot of the two measures is given in Figure 1. From this plot two individuals were selected as examples of individuals for whom the two approaches yielded very different classifications of complexity. (See Table 2). As can be seen in Table 2, subject A rated six persons (columns) in a highly differentiated and complex fashion and rated the remaining persons in an undifferentiated or simple fashion. This yielded a very low and therefore cognitively simple Bieri's d of 168, since most columns contained similar ratings. However, the overall pattern of ratings within each row (trait) were sufficiently varied so that rating on any one row could not be predicted. This resulted in a very low first root (1.91) suggesting cognitive complexity.

Subject B's grid, on the other hand, is an instance of exactly the opposite. Ratings for each person were made in a very diffuse fashion, yielding a very high and therefore cognitively complex d of 353. However, the overall pattern of ratings within each row was sufficiently similar, so that the first root was quite large, (5.69) suggesting cognitive simplicity.

Correlations with Other Instruments

The second step in comparing the alternate scoring methods with Bieri's d consisted of a comparison of the correlations of the scores from each of the scoring methods with three other instruments thought to be related to cognitive complexity. These instruments were the Group Embedded Figures Test, the Paragraph Completion Test, and a student evaluation questionnaire. For an indepth discussion of the conceptual relationships between these measures and the Rep Test please see Trabin & Doyle (1981).

The Group Embedded Figures Test (GEFT) is a ten-minute timed test consisting of 18 complex designs from which the subject must disembed a specified geometric figure; lower scores (less disembedded designs) indicate greater field dependence. Field dependence/independence is a bipolar cognitive style defined as "the extent to which a person is able to deal with a part of a field separately from the field as a whole, or the extent to which he/she is able to disembed items from an organized context." (Witkin, 1976, pp. 41-42).

The Paragraph Completion Test (PCT) is used to measure integrative complexity and requires that the subject elaborate six sentence stems into paragraphs (i.e., "When I am criticized...", "When I am told what to do...", "When I am not sure..."). This is essentially a cognitive developmental measure, with development conceptualized as movement from a concrete to an increasingly abstract conceptual system.

The student evaluation of instruction questionnaire used in this study included 12 items which pertained directly to instructor abilities. Each item was in a six-point Likert format with anchors ranging from "strongly agree" to "strongly disagree." Differentiated student ratings was measured by taking the standard deviation of the 12 ratings for each subject.

Table 3 shows the correlations of these three measures with the three approaches to scoring the Rep Test. As can be seen, the correlations between these measures are only significant in two cases, and the first characteristic root failed to correlate with any of the scores. The low, but significant correlation between the standard deviation of student ratings and Bieri's d is a replication of earlier findings (Wright & Richardson, 1979).

Conclusions

Cognitive complexity has been conceptualized differently by researchers using different measurement instruments and scoring techniques. It has been defined as the multidimensionality of social judgment by those researchers using the modified Rep Test and scored employing Bieri's d. This definition and measure of cognitive complexity have received more attention than any other approach to cognitive complexity in the psychological literature. However, as this study has demonstrated, Bieri's d is not a measure of multidimensionality of social judgment, and the correlation between d and the measure of multidimensionality proposed here was, for practical purposes, zero.

As was expected, Sierri's d did correlate moderately with the average standard deviation of responses within columns. These findings lead to the conclusion that Bieri's d is probably better considered a crude measure of dispersion within columns, or, possibly, response style.

All measures of multidimensionality failed to show correlations with three other measures which also attempt to assess complexity or analytic ability. The slight exception to this finding were the low, but significant correlations of d and the average standard deviation approach with the standard deviation of student ratings. This is surprising, given the similar definitions of analytical/complex behavior outlined in the theoretical rationale for the instruments. This finding further reinforces the conclusion that Bieri's d is probably not related to multidimensionality.

Limitations

The limitations of scaling in this study are quite apparent. First, no estimates of the reliability of the alternate Rep Test scales were attempted. The degree of generalizability of an eigenvalue based on a single individual's ratings is simply not known, due to the violations of independence, the limited ratio of variables to "observations" and the restriction of range inherent in the Rep Test itself.

Implications for Future Research

Given these findings, future directions for fruitful research in this area might investigate the comparative reliability, validity, and usefulness of each perspective on cognitive complexity. For instance, future research might evaluate the degree to which methods outlined here for scoring the modified Rep Test are related to previous cognitive complexity measures on Kelly's original Rep Test. Also, the comparative efficacy of these different operationalizations may be studied with respect to prediction of various types of social perception.
dilemmas. Finally, other approaches to operationalizing cognitive complexity should be explored. In this vein, a multidimensional scaling approach may show promise as a model which meets the ordinal assumptions of the data, and provides fit measures to identify subjects who completed the instrument randomly or failed to understand the directions. Studies along the above-mentioned directions will help to elaborate the nomological net for cognitive complexity. It is hoped that this study will have been a contribution toward increasing the precision of definitions and operationalizations of cognitive complexity.

Acknowledgements

We wish to acknowledge the assistance of Kenneth Doyle in the design and collection of data for this study. We also wish to acknowledge the assistance of Auke Tellegen in suggesting alternative scoring approaches.

References


## Cognitive Complexity Scoring Algorithms

<table>
<thead>
<tr>
<th></th>
<th>Group Embedded Figures Test</th>
<th>Paragraph Completion Test</th>
<th>Differentiatedness of Student Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bieri's &quot;d&quot;</td>
<td>-.01</td>
<td>-.04</td>
<td>.22**</td>
</tr>
<tr>
<td>standard deviation</td>
<td>-.03</td>
<td>-.07</td>
<td>.22**</td>
</tr>
<tr>
<td>first Eigenvalue</td>
<td>-.06</td>
<td>.00</td>
<td>.04</td>
</tr>
</tbody>
</table>

* Pairwise deletion resulted in N's between 257 and 320 for the different correlations.

** p < .005