APPENDIX D

EXPLORATORY FACTOR ANALYSIS

Factor analysis is a data reduction technique that reduces the number of variables used in an analysis by creating new variables (called factors) that combine redundancy in the data. The factor analysis conducted for this report reduced the number of variables from ten to four factors. The underlying theory behind factor analysis is that a set of multivariate observations stem from a lesser number of underlying factors. For example, data on students' test scores in math, science, art, and literature might reflect two underlying factors: one for "technical ability" and one for "creativeness." A student's "technical ability" will probably be better reflected in math and science test scores, while a student's "creativity" should be reflected in art and literature test scores. The factors are meant to reflect some underlying abstract dimension or concept. The measured variables are imperfect measures of those dimensions or concepts.

A factor analysis looks for trends in correlations among the variables. Given that the factors are unmeasured, it is necessary to use numerical algorithms to solve a factor analysis. The first step in a factor analysis is to determine the number of relevant factors. Many algorithms used to solve factor analyses have methods of determining an appropriate number of factors, but it is also possible to specify (fix) a number of factors. For the analysis in this report, we allowed the algorithm to determine the number of factors.

The output from the factor analysis will generate a table that relates each variable to each factor and assigns a numerical value between -1 and 1 to each relationship. The numerical values are referred to as factor loadings and reflect the strength of relationship between the factors and the variables. Variables that are closely related to one another should all load highly on the same factor. This is the essence of factor analysis: combining redundant variation in the data.

The factors generated in a factor analysis do not have a straightforward interpretation. In fact, it is up to the researcher to determine how a factor should be named. For the most part, theoretical considerations can guide in naming a factor. Nevertheless, developing appropriate names for factors is an important aspect of factor analysis.

Once a factor analysis has been performed, a mathematical operation called *rotation* is performed. The purpose of rotation is to make each factor distinct (in terms of factor loadings) from the other factors. Most raw factor loadings require rotation. The term rotation stems from the technique involved: the axes are literally rotated around the score to generate new axes and thus new factor scores. From a mathematical perspective, this transformation is justified since the factor loadings are only unique up to a multiplicative constant. Thus, rotation need only preserve the order of the loadings to be consistent.

Scoring is the process of generating values for the factors for each observation in the data. For example, a factor analysis that reduces a set of 20 variables to six factors might be based on 1,000 observations on those 20 variables. The factor analysis only generates 120 factor loadings (20 variables × six factors). Although each observation has a value for each variable, none of the observations has a value for the six factors (at this point). Scoring assigns a value to each observation for each factor. Once again, because the factors are unobserved, it is necessary to use numerical algorithms to solve the equations used to score the factors.