Common Factor Analysis: The Early Years

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

Common Factor Analysis: The Early Years

Introduction

- In this module, we discuss factor analysis, an extremely popular data analytic technique that dates back to the beginning of the 20th century.
- Exploratory Factor Analysis (EFA) is the historical precursor to confirmatory factor analysis and structural equation modeling.
- Major books have been written about factor analysis, and focus of this module is on the key algebraic properties of the factor analysis *model*.

Spearman's Theory of g

Measuring the Unmeasurable

- In 1904, Charles Spearman, a British psychologist, proposed his "single factor" theory of intelligence.
- Spearman sought to explain the relationships among various measures of mental ability by means of a single (underlying) ability, which he called general intelligence, or "g."
- Spearman believed that g was the common thread underlying performance on all tests of mental ability.
- In his view, each mental test tapped a general mental ability and a specific ability.
- Spearman's g was a "latent" variable, in the sense that there did not exist independent operations and criteria for measuring it.
- Rather, it was defined only in terms of the equations of the factor analysis model.

Spearman's Theory of g

Measuring the Unmeasurable

- Spearman postulated that g existed, even though it was only evidenced indirectly by the battery of mental ability tests.

How could g be uncovered?

Spearman's Theory of g

Measuring the Unmeasurable

- The *possibility* of the existence of a g could be tested.
- Suppose g exists, and explains the correlations among mental ability tests in the "partial correlation sense."
- Then, if we could measure g, and partial it out of the mental ability tests using linear regression, the partial covariances should all become zero.
- Suppose the observed variables are gathered in a random vector $\mathbf{y}$.
- Then, if the general intelligence factor g explained the correlations in the variables in $\mathbf{y}$ in the partial correlation sense, the residual covariance matrix for the variables in $\mathbf{y}$, with g partialled out, should be a diagonal matrix.
- But the question remained — since g can't be measured directly, how could one test this empirically?

Spearman's Theory of g

Measuring the Unmeasurable

- Let's trace the steps carefully.
- Since the latent variable g is never actually observed, one might, with no loss of generality, imagine its variance to be 1.
- One then asks, "Suppose g existed and we had a way of measuring it directly. What empirical evidence would support (or falsify) Spearman's hypothesis?"

Spearman's Theory of g

Measuring the Unmeasurable

- Spearman deduced that the existence of a g could be verified by showing that a vector of regression weights $\mathbf{f}$ exists such that $\mathbf{\Sigma_{yy}} - \mathbf{ff}'$ is diagonal.
- Clearly, for a given $p \times p$ covariance matrix $\mathbf{\Sigma_{yy}}$ with $p > 2$, there may not be any $\mathbf{f}$ such that $\mathbf{\Sigma_{yy}} - \mathbf{ff}'$ is diagonal, and so Spearman's model was falsifiable.
- Determining exactly *how* the model could be falsified was a significant achievement, and we will examine it in a companion lecture.

Early Optimism

Measuring the Unmeasurable

- Once Spearman realized that his model could be tested, and that preliminary results seemed to support it, he spent a number of years gathering data on mental ability tests in the hope that it would verify his model.
- He hoped that a number of benefits would ensue from fitting the common factor model (with a single common factor) to a set of mental ability tests.
  1. by fitting the common factor model and determining $\mathbf{f}$, the factor loadings, he hoped to discover which ability tests loaded highest on general intelligence.
  2. By obtaining the sample equivalent of the random variable g, i.e., the vector of observed intelligence factor scores, he hoped to be able to obtain a pure measure of intelligence for each individual.
- This intelligence score could, ultimately, be registered for each person, and help determine that person's position in the society.
- By 1927, his work had progressed, empirical support had been gathered, and he embarked on an American lecture tour to promote his book, *The Abilities of Man*.

Conceptual Problems and Controversies

Measuring the Unmeasurable

- There were a number of complications that sidetracked Spearman in his ambitions.
  1. *The Problem of Sampling Variability.*
  2. *The Spearman-Thomson Controversy.*
  3. *Competition from Multiple Factor Models.* Other researchers believed there was more than one fundamental factor of mental ability. In terms of the ability to fit data, their models had a built-in advantage. L.L. Thurstone, at the University of Chicago, wrote two very influential books, *The Vectors of Mind* (1936) and *Multiple Factor Analysis* (1947), in which he promoted his multiple factor model as an improvement on Spearman's approach.
  4. *The Indeterminacy Problems.* Things weren't as mathematically straightforward as they seemed at first glance. When Spearman arrived at Harvard on the first stop of his lecture tour, he encountered a famous American mathematician, E.B. Wilson, who challenged him regarding the mathematics of his model in ways he had not anticipated.

Conceptual Problems and Controversies

Measuring the Unmeasurable

- Ultimately, of course, as computers grew in power, factor analysis became a popular tool for analyzing data in a variety of fields.
- One reason for its popularity was that the model could be justified from a number of different rationales (Steiger, 1994)
  1. *The Partial Correlation Rationale.*
  2. *The Random Noise Rationale.*
  3. *The True Score Rationale.*
  4. *The Data Reduction Rationale.*
- By the late 1960's, factor analyses that took a year to complete just a decade before could be finished in an afternoon.
- Interestingly, in the 1970's, there was a reawakening of interest in the conceptual issues that sidetracked Spearman 40 years before.

Conceptual Problems and Controversies

- In the next module, we examine how Spearman sought to test the factor model with data by examining the pattern of correlations among the observed variables.
- Spearman's *algebraic* approach, though now largely ignored in favor of more statistically oriented approaches, has much to teach us, and has been periodically revisited by factor analysts.