ABSTRACT

Title: ACCURACY AND CONSISTENCY IN DISCOVERING DIMENSIONALITY BY CORRELATION CONSTRAINT ANALYSIS AND COMMON FACTOR ANALYSIS

Rochelle Elaine Tractenberg, Ph.D., 2009

Directed By: Professor Gregory R. Hancock, Department of Measurement, Statistics and Evaluation

An important application of multivariate analysis is the estimation of the underlying dimensions of an instrument or set of variables. Estimation of dimensions is often pursued with the objective of finding the single factor or dimension to which each observed variable belongs or by which it is most strongly influenced. This can involve estimating the loadings of observed variables on a pre-specified number of factors, achieved by common factor analysis (CFA) of the covariance or correlational structure of the observed variables. Another method, correlation constraint analysis (CCA), operates on the determinants of all 2x2 submatrices of the covariance matrix of the variables. CCA software also determines if partialling out the effects of any observed variable affects observed correlations, the only exploratory method to specifically rule out (or identify) observed variables as being the cause of correlations among observed variables. CFA estimates the strengths of associations between factors, hypothesized to underlie or cause observed correlations, and the observed

variables; CCA does not estimate factor loadings but can uncover mathematical evidence of the causal relationships hypothesized between factors and observed variables. These are philosophically and analytically diverse methods for estimating the dimensionality of a set of variables, and each can be useful in understanding the simple structure in multivariate data. This dissertation studied the performances of these methods at uncovering the dimensionality of simulated data under conditions of varying sample size and model complexity, the presence of a weak factor, and correlated vs. independent factors. CCA was sensitive (performed significantly worse) when these conditions were present in terms of omitting more factors, and omitting and mis-assigning more indicators. CFA was also found to be sensitive to all but one condition (whether factors were correlated or not) in terms of omitting factors; it was sensitive to all conditions in terms of omitting and mis-assigning indicators, and it also found extra factors depending on the number of factors in the population, the purity of factors and the presence of a weak factor. This is the first study of CCA in data with these specific features of complexity, which are common in multivariate data.

ACCURACY AND CONSISTENCY IN DISCOVERING DIMENSIONALITY BY CORRELATION CONSTRAINT ANALYSIS AND COMMON FACTOR ANALYSIS

By

Rochelle Elaine Tractenberg

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Professor Gregory R. Hancock, Chair
Professor Robert J. Mislevy
Professor Robert W. Lissitz
Professor George Macready
Professor Kevin E. O'Grady

Dedication

ded·i·ca·tion. According to the Merriam-Webster Dictionary definition, this noun dates back to the 14th century. Among the definitions are such diverse elements as: 1. An act or rite of dedicating to a divine being or to a sacred use. 2. A devoting or setting aside for a particular purpose. 3. A name and often a message prefixed to a literary, musical, or artistic production in tribute to a person or cause. 4. Self-sacrificing devotion. 5. A ceremony to mark the official completion or opening of something.

Too many individuals, only a few of whom are listed in the acknowledgements, merit tributes with respect to the official completion of my doctoral education, which this work (admittedly neither literary nor artistic) represents. By official and completion, I mean that I am finally done with doctoral work- after two PhDs and a doctoral-level certificate. It is *over*. David Liebke and Seven and Emma Tractenberg-Liebke are the ones who demonstrated self-sacrificing devotion to my completion of this work, with their tremendous collective patience and perhaps more importantly, their persistence in requiring play breaks.

Acknowledgements

This work could not have been possible if not for the support of my mentors, Paul Aisen and Gregory Hancock. The work was supported by NIH Grant K01-AG027172 from the NIA and, very early on, by a Research Starters Grant from the Pharmaceutical Research and Manufacturers of America Foundation. The Department of Neurology at Georgetown University Medical Center was very generous in terms of time and support, especially the Chair, Dr. Edward Healton. David Liebke and Futoshi Yumoto provided absolutely invaluable assistance with computer programming and automation features of the study; David made Tetrad interesting to me in the first place, and made it useable in the last place. I hope to be even half as helpful and supportive of their doctoral work as they have been of mine!

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and overview of multivariate methods for uncovering simple structure

The psychometric characteristics of many instruments, as well as collections of variables, used in social science and other types of research are frequently established by estimating (or as a function of) the underlying dimensions of a scale, instrument, or set of variables. Multivariate analysis is used to estimate the number of dimensions, typically with the objective of finding the single factor or dimension to which each observed variable belongs or by which it is most strongly influenced. When more than one dimension is present, investigators seek simple structure (Thurstone, 1935; Yates 1987), in which each observed variable is associated (most strongly) with a single underlying or latent variable. This “between-variables” multidimensionality (Wang, Wilson, & Adams, 1997) describes instruments or sets of variables that are most often studied in real-world situations where a single factor is either insufficient to capture the complexity of the data or where the data itself was intended to represent multiple dimensions. This dissertation compared and contrasted two different methods for estimating the dimensionality (i.e., number of latent variables and their constituent indicators) in a collection of variables, under the assumption that at least approximately, each observed variable loads on exactly one latent variable.

1.1 Estimating dimensions

1.1.1 The correct number of dimensions is critical

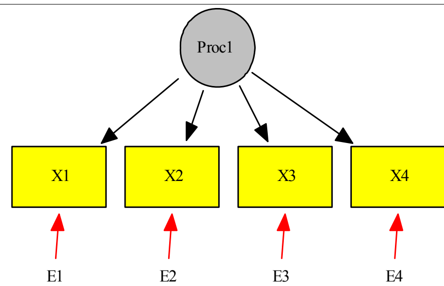
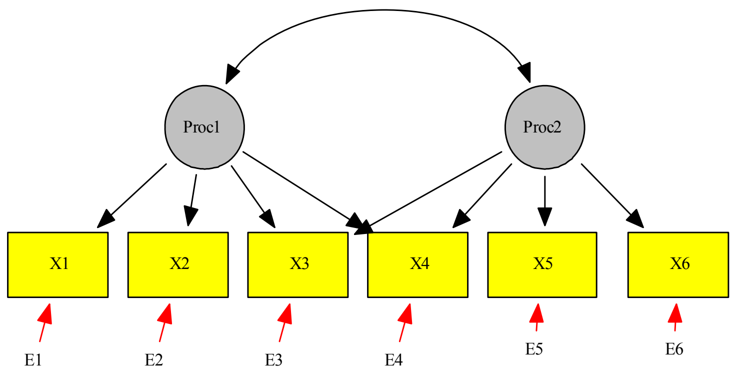
Wang, Wilson, and Adams (1997) described measurement in the testing context within which psychometric modeling (Rasch models) are the basis for utilizing the dimensionality –specifically, the unidimensionality – of a set of test items to make inferences about some construct (the purpose of the test). Outside of testing and item response theory, collections of questions or items are often used to assess constructs that are theorized to be multidimensional. Very often, a set of responses is obtained and the relationships among the questions eliciting the responses are the topic of study.

If the underlying structure of an instrument is incorrectly determined to be unidimensional, unexpected error and bias may result although an instrument itself may actually be valid (i.e., contain appropriate items, reflect the underlying construct, correspond to benchmarks/criteria, and support inferences about the target population) and reliable (providing the same information about individuals over repeated administrations). For example, psychological and neuropsychological assessments can be used/useful without any formal appreciation or estimation of their underlying dimensionality.

However, to modify instruments, or to document their utility in different contexts (e.g., clinical trials) or populations (e.g., to support inferences about different groups) than were originally intended when the instrument was developed, exploration of dimensionality is sometimes pursued –often in an exploratory manner. In his description of methods for determining the optimal number of factors to extract in exploratory analyses, O'Connor (2000) noted, "under-extraction compresses variables into a small factor space, resulting in a loss of important information, a neglect of potentially important factors, a distorted fusing of two or more factors, and an increase in error in the loadings." (p. 396). In their brief review of the literature describing and discussing the effects of over- and under-extraction of factors, Velicer, Eaton, and Fava (2000) argued that incorrectly identifying (or estimating) the number of factors, or components, will lead to problems with solution variability across samples (p. 42) –with scientific implications for factorial and configural invariance across samples. Velicer et al. (2000) agreed with O'Connor (2000) in the conclusion that under-extraction is the worse, and more consistently damaging, result of mis-specification of the true dimensionality of the variable/score set. Further, Kano (2007) noted that "serious bias can be created by analysis with a model containing inconsistent variables…" (p. 65), meaning that observed variables must be included in factor analysis (and other latent variable models) that are consistent with the model itself.

If instruments are erroneously modeled as unidimensional – whether dimensionality is mis-estimated or items that do not fit the dimension of interest are included, then the statistical models based on that incorrectly-conceptualized instrument will not be useful. This bias can lead to inconsistent findings, inappropriately low power in experiments, and other difficulties in research. Some practical implications arising from incorrect estimates of dimensions are exemplified in Figure 1.

Figure 1. Incorrect dimensionality: impact on future science

<table>
<tr>
<td>
</td>
<td>One factor (process, Proc1) is assumed to underlie the observed correlations between the variables X1-X6. The four indicators are deemed sufficient indicators of the single process and interventions are developed targeting that single dimension.</td>
</tr>
<tr>
<td>
</td>
<td>Two factors (processes, Proc1, Proc2) are assumed. The four indicators that were previously deemed sufficient indicators are now augmented by two additional indicators (X5, X6) which would not have been considered-or collected - under the assumption of one process, and interventions are developed targeting both dimensions.</td>
</tr>
</table>

When a set of variables is incorrectly 'determined' to have a single underlying (causal) factor, this can limit the scope and direction of future research. In the lower panel of Figure 1, two indicators (X3, X4) are causally associated with both of two latent factors; in addition two other indicators (X5, X6) that would not be collected or sought, because they are only associated with the second latent variable, are shown in the lower panel. This example reflects the situation where the model is theoretically –not just statistically – misspecified.

Consider attention deficit-hyperactivity disorder (ADHD) as being represented (without any appeal to clinical reality) in Figure 1. This totally fictitious example then explores the impact of mistakenly identifying a single dimension to

ADHD. Assuming that hyperactivity is the ‘salient’ observable behaviour, i.e., the single factor or dimension that is identified (“Proc 1”), treatments or strategies will focus on indicators or symptoms of hyperactivity. Since attention, and/or executive control of attention, may be more likely to be responsible for observed classroom (or workplace) difficulties encountered by an individual with ADHD (see, e.g., the role of encoding in comprehension described by Brown and Craik, 2000), interventions addressing only hyperactivity (or a single ‘causal’ mechanism) could be less likely to improve classroom outcomes. Since variables in Figure 1 representing attention deficit (“Proc 2”) specifically (X5, X6) would not be considered, or collected, if the only dimension was hyperactivity (e.g., top panel), the science of understanding ADHD, and its treatment(s), could be slowed.

The impacts of having estimated the incorrect number of dimensions in this example would be misdirected interventions, incorrectly limited variable collection (which could hamper future efforts to understand the disorder and to develop effective interventions), and greater difficulty overall in validating, or ameliorating, the effects of ADHD in the indicators (X1-X6). These practical implications of the mis-specified model are separate from any statistical implications.

1.1.2 Methods for estimating the number of dimensions

Principal components analysis is a method for reducing the dimensionality of a set of variables without considering how many dimensions are actually present. That is, data reduction (by principal components analysis) takes a set of p variables and converts it into a collection of m ($m<p$) orthogonal variables that are composites of the original variables. Changing the dimensionality, or mapping a set of p variables

onto a smaller dimensional space, is not the same (neither theoretically nor mathematically, Widaman 2007) as determining the dimensionality of the original *p* variables.

A common method to determine the dimensionality of a set of items (scale, instrument) or variables for a given population is exploratory factor analysis (for an alternative, see Zhang and Stout, 1999 who describe a novel approach to determine if data have only one dimension and describe other methods for assessing unidimensionality specifically). One way that exploratory factor analysis can proceed is common factor analysis. In common factor analysis (CFA), the investigator specifies the number of factors believed to (causally) underlie the observed variables, and the algorithm produces the factor loadings for observed variables that are most compatible with both that number of factors and the correlations among the observed variables (Widaman, 2007).

Algorithms most often used in factor analysis include maximum likelihood (ML, e.g., Bartholomew & Knott, 1999) and two-stage least squares (2SLS, Bollen, 1996) for estimating associations between observed and latent variables (“loadings”). Briggs and MacCallum (2003) showed that, in the presence of weak factors, ordinary least squares (OLS) performs better than ML due to the fact that ML seeks greater precision for estimates involving larger correlations when smaller correlations (“weak” factors) are also present. OLS estimation is infrequently utilized in analyses involving latent variables, but the Comprehensive Exploratory Factor Analysis Package (CEFA, Browne, et al. 2008), used by Briggs and MacCallum to apply OLS to their latent variable extraction/estimation problems (and to compare the results with those obtained using ML; SPSS (SPSS Inc., Chicago, Ill.) and Mplus (Muthen & Muthen, 2008) also offer an unweighted (ULS, Mplus) or ordinary (OLS, SPSS) least squares estimation options, albeit without the possibility of estimating standard errors for the estimated loadings. As will be described in Chapter 2, OLS estimation will be used in this study.

Importantly, a great deal of the literature describing exploratory factor analysis (FA) is aimed at demonstrating how incorrect results – in terms of factor loadings - can be obtained when principal components analysis, rather than common factor analysis, is used (e.g., using Kaiser's "little jiffy", described in Preacher and MacCallum, 2003 and Widaman, 2007). While interpretation of factors, and possibly identification of the "correct" number of factors underlying the data set, may proceed on the basis of loadings of items/scores on the extracted factors, many exploratory factor analyses are carried out by investigators who are actually concerned with uncovering, or discovering, the latent structure; interest in loadings in that context is only in terms of whether they can clarify the interpretation/interpretability of the factor structure. However, as noted by Silva, et al (2006) and Spirtes, et al. (2000), the latent structure is not directly available from exploratory factor analysis methods, irrespective of the extraction method. Unidimensional latent structure is suggested by the presence of so-called vanishing tetrad differences together with non-vanishing correlations when the observed variables are partialled out. Latent *structure* is not directly supported by reductions in rank of a matrix, the maximization of variance explained in a system of variables by a smaller set of factors, or rotation to simple structure. These features of exploratory factor analysis are detailed further in Chapter 2.

In short, factor loadings are often used to interpret the factors implied by the loadings; rotation facilitates this (although Widaman, 2007, noted that rotation "destroys the conditional variance maximization property" that a reduced rank matrix can provide, p. 187), but loadings cannot directly support conclusions about latent structure, i.e., the causal relationship between the latent factor and the manifest variables that, together, represent a measurement model. Although Thurstone (1937; 1940; 1947) advocated using factor loadings to assist interpretation and selection, or identification of the "appropriate" solution, these estimates themselves do not provide information about latent structure; they only represent the interrelations of the indicators given the current solution.

Another body of literature is devoted to methods for determining how many dimensions there are in the data. This literature sometimes clearly pertains to common factor analysis, sometimes clearly pertains to principal components analysis (PCA), and sometimes is not clear to which analytic method a given article/book pertains. The use, and utility, of factor analytic methods can be obscured by diverse representations across these reference materials. The confusion may be a reason why users of these methods tend to rely on the default settings of software (Osborne, Costello, & Kellow, 2008).

Bartholomew and Knott (1999) suggested that, since statisticians may tend to prefer PCA over "factor analysis" (FA, also called common factor analysis), and because some software is perhaps less explicit, or less theoretically oriented, than it could be, PCA and FA are often confused. However, Bartholomew and Knott also stated that apart from reducing the dimensionality of data, latent variables are useful when the construct of interest cannot be measured, or cannot be measured directly. In that context they note that when observed variables (e.g., items or questions) are collected together, latent variable methods are used to "…extract what is common to them." (p. 2). They further emphasized this point in the formulation of the axiom of conditional independence (pp. 4-5); "A key part of our analysis is directed to discovering the smallest q…" that conditional independence will adequately fit. Finding the smallest number of latent variables is not the same as finding the latent variables and measurement models that fit both the observed and the theoretical information best.

Afifi, Clark, and May (2004) noted that "Ideally, the number of factors expected is known in advance" and imply that the purpose of factor analysis is to obtain factor loadings that explicate relationships among observed variables (p. 392). This purpose – for common factor analysis, or CFA, is contrasted with Harman's (1976) and Widaman's (2007) characterizations of the two main goals (maximize variance explained with composites or minimize differences between model-implied and observed correlations), but optimizing the factor loadings is a key component in the estimation procedures of all exploratory factor analytic methods. Thus, "exploratory" could be exchanged for "loadings-estimating" in characterizing these approaches.

In contrast to Afifi, et al. (2004), Bandalos (contributing to Stevens, 1996) stated that "(t)he purpose of exploratory factor analysis is to identify the factor structure, or model for a set of variables. This often involves determining how many factors exist, as well as the factor loadings.” (p. 389). Similarly, Pett, Lackey, and Sullivan (2003) suggested that the purpose of factor analytic methods is, “… through data reduction, to group a smaller set of these <observed> variables into dimensions or factors that have common characteristics” (p. 2). This description is more or less correct in that exploratory FA usually does reduce the number of observed variables to a smaller set of variables that have common characteristics; however, reducing the dimensionality of a set of variables is not the same as uncovering the latent structure or measurement model underlying the correlations between the indicators.

Although PCA approaches are common, and commonly advised for new exploratory FA users (or as a first pass to estimate the number of factors, Velicer, et al. 2000), using exploratory FA to identify factor structure, estimate loadings, and reduce the dimensionality of a data set confounds the two methodological approaches to exploratory analyses that Widaman (2007) carefully distinguished: identifying common factors or explaining maximum variance with a concise number of principal components (or factors). Possibly contributing to the confusion is the reliance on loadings, from a definition from the early 1960s, that a factor is a construct which is operationally defined by the loadings on it of observed variables (Royce, 1963; recapitulated in Kline, 1994, Ch. 1). Thurstone (1937; 1940; 1947) advocated that loadings (after rotation) should be used as the evidence for experimentally testable hypotheses about the underlying factors- but he also advocated careful experimental design and theoretically informed tests/batteries as the basis of exploratory factor analysis.

1.1.3 The purpose of factor analysis

As exploratory FA has become more widely used and more quickly obtained, the "purpose of factor analysis", has devolved from Thurstone's original ideal: "…to know whether …measures are related by some underlying order which will simplify our comprehension of the whole set of measures…" (Thurstone, 1940, p. 216). The original perspectives on finding common factors (evidence of one common factor <Spearman>) and evidence of experimentally-testable psychological traits <Thurstone>, have shifted: as Hoyle and Duvall (2004) noted, "(t)he aim of factor analysis is to describe the associations among a potentially large number of observed variables, or indicators, using a relatively small number of factors" (p. 301). Today's researcher may be using exploratory FA "…when the researcher does not know how many factors are necessary to explain the interrelationships among a set of characteristics, indicators, or items" (Pett, et al. 2003, p. 3). However, exploratory FA is not the correct choice if the underlying structure – measurement model- for a set of observed variables is desired. Bollen (1989) noted that within a factor analytic framework, all observed variables are assumed to be caused by the latent factor (p. 7), representing a measurement model (a latent variable and the indicators it is causally related to). In some cases, the latent variable (e.g., socio-economic status) is caused by, and does not cause, the indicators to vary; exploratory FA will not distinguish these two types of models.

Velicer, et al. (2000) surveyed the literature, and did an empirical study, regarding the use of a variety of techniques for identifying the 'correct' number of dimensions or factors underlying a data/variable set. Importantly, they noted that there was (as of 2000) no reliable method for accurately estimating the number of factors – although two methods (Horn's Parallel Analysis, Horn, 1965 and Velicer's Minimum Average Partials, Velicer, 1976) have been repeatedly demonstrated to be accurate in the recovery of the correct number of principal components. Velicer et al. recommend using either of these methods for principal components analysis, which is a fundamentally different method than common factor analysis (CFA, see, e.g.., Widaman, 2007), but they further recommend using a principal components analysis approach to "guide" the estimation of the number of factors underlying one's data (p. 68).

1.1.4 Estimating dimensionality by correlation constraint analysis

A far less studied and less well-known method for estimating the dimensionality of a set of variables or items is correlation constraint analysis (CCA). CCA proceeds in a manner wholly distinct from that of CFA and any other factor-analytic approach. CCA is actually the forerunner of modern factor analysis; it was the method developed by Spearman (1904; 1927) to support his 'two factor' theory (where one of the two factors is common to all items, while the second factor is unique to each item, in the analysis) of intelligence. Unlike common factor analysis (and any other dimension-estimation or rank-reducing method), CCA (as implemented by TETRAD, the CCA software) operates on the determinants of the 2x2 submatrices of the covariance matrix. Software using CCA also determines if partialling out the effects of any observed variable affects the correlations that are observed- thus, CCA is the only exploratory method to specifically rule out any observed variable as being the cause of correlations among observed variables. This can be contrasted with observing a Heywood case in a CFA solution, suggesting that one of the observed variables could be the cause of other observed variables (but actually indicating a boundary-value or negative variance): the CCA approach (by TETRAD) specifically tests each observed variable by partial correlation to determine if removing its influence statistically significantly alters the correlation value before the partialling.

The specific details of how CFA and CCA are achieved, and how they differ, are presented in Chapter 2. Overall, as noted by Bollen (1989, p. 4), CFA is useful to test whether causal assumptions represented in a hypothesized model are statistically consistent with the observed data (covariance or correlation matrices), while CCA is useful to uncover evidence of causal relationships. These are philosophically diverse methods for estimating the "true" dimensionality of a set of variables or items. An additional point of divergence is that CFA seeks to estimate the strengths of associations between observed and hypothesized latent variables, while CCA seeks to estimate the strengths of associations between observed variables –but only insofar as to determine if there is sufficiently strong association to conclude, and in the pattern consistent with, the existence of common causal factors. Based on these estimations, CCA determines whether algebraically invariant relationships between correlations hold (Bollen 1990; Bollen & Ting, 1993; Bollen & Ting, 2000; Spirtes, et al. 2000; Silva, et al. 2006; see also Drton, et al., 2007). The software utilized within this study follows that with estimations of partial correlations – where the effects of each observed variable are partialled out of the estimated correlations between every other pair of observed variables.

Critically, when the algebraically invariant relationships among sets of four observed variables hold and the partial correlations do not change when the influence of other observed variables is removed, conclusions about the presence of latent common causal variables are supported. When one or neither of these features is observed, other conclusions are supported. Thus CCA results are based on the estimated correlations, and partial correlations, within the observed variables. There is no estimation that optimizes the discrepancy between model-implied and observed covariability. These features of the two methods (CFA, CCA), which were compared in this dissertation, are described more fully in Chapter 2.

1.2 Overview of Dissertation

1.2.1 The research question: Is CCA as good as CFA, or better, at uncovering dimensionality?

One of the motivating issues for this dissertation is that many social science applications of factor analysis include higher order factors, and structural models connecting latent variables (or, rather, their measurement models, see Bollen, 1989). Factor analysis focuses on first order measurement models (and principal components analysis), although correlations between factors can be estimated in second stage factor analysis (i.e., after the first solution is obtained, the analysis is repeated on the factors, to extract factors relating, or estimate strengths of associations between, factors). By contrast, CCA can be applied to uncover measurement models and (subsequently, within the software TETRAD (http://www.phil.cmu.edu/projects/tetrad_download)) the presence of structural relations among latent variables. This dissertation research is a first step towards introducing CCA to a wider audience, by comparing the consistency and accuracy of CCA and CFA in uncovering the dimensionality in a set of variables simulated to have mainly simple structure (but with some variability) and as having first-order causes.

Very few researchers are aware of CCA as an alternative (method or philosophical position), and the relative performances of CCA and CFA have not yet been compared along the dimensions this dissertation proposes. However, given the importance of identifying the correct number of latent variables, or the 'dimensionality' of a group of variables/scores, and the specific strengths and limitations that CFA and CCA bring to this problem, this dissertation proposes to study and document the performance of CCA at uncovering the latent structure of simulated data. CCA will be tested under conditions of varying sample size, model complexity, the presence of weak common factors, and correlated vs. independent latent variables. This research will determine if the performance of CCA is just as, or more, factorially invariant than that of common factor analysis (CFA) under these conditions. Factorial invariance represents the generalizability of the test results across persons and time (see, e.g., Millsap & Meredith, 2007), as well as supporting response process modeling validity evidence (Boorsbom, 2005). Previous studies of CCA (Silva, 2005; Silva, et al. 2006; Spirtes, et al. 2000) have not focused on the consistency of solutions over multiple iterations, but have instead reported the average number of omission (paths incorrectly left out) or commission (paths incorrectly put in) errors across 10 (Silva, 2005; Silva, et al. 2006) or 20 (Spirtes, et al. 2000) trials. Further, these simulations have all employed maximum likelihood estimation (ML or MLE), which may be more subject to erroneous results when weaker path loadings, relative to stronger ones, are present in the true model (Briggs & MacCallum, 2003).

1.2.2 Organization of the dissertation

This dissertation is the first study of how CCA compares to CFA by OLS in identifying the dimensionality of data varying in terms of specific features of complexity often encountered by social scientists, and is also the first study to explore accuracy and consistency (invariance), and whether these two characteristics are sample size dependent, for CCA. The research will answer the questions of whether CCA and CFA perform at similar levels in terms of accuracy and consistency over repeated trials, and whether this performance will be affected by features of sample size, model complexity, the presence of weak common factors, and correlated vs. independent latent variables. Chapter 2 outlines the dimensionality problem more explicitly with a literature review, and articulates how common factor and correlation constraint analyses differ. The research methods are given in Chapter 3. Results of the simulation study will be given in Chapter 4 and discussion, conclusions and recommendations for uncovering dimensionality will make up Chapter 5.

Chapter 2: Background and Rationale for Dissertation

This chapter describes the problem of estimating dimensionality over the history of the factor analysis methodology development (over roughly 100 years). This review includes the literature describing how dimensionality is estimated by current, widely available software. Estimation of dimensionality by correlation constraint analysis is also described.

2.1 The problem: uncovering dimensionality vs. consolidating variables

This study targeted situations where the investigator is interested in learning the number of latent variables that are responsible for (caused) the observed correlations among a set of variables or scores. Estimating the dimensionality of a set of variables is not always a search for the underlying causal mechanism; for example, Bentler (2007) described the situation where the investigator wishes to estimate the reliability of subsets of items in an instrument, and evidence of unidimensionality –to support reliability estimation, but not a causal influence – is sought (see also Zhang & Stout, 1999). In some cases, this is labeled an “exploratory” research problem, that is, no hypothesis drives the analysis, the data do. The exploratory approach to estimating dimensionality is to see how many latent variables or factors the correlation or covariance structure supports. This can (but need not) be contrasted with “confirmatory” approaches, where a hypothesis (e.g., “these observed variables are caused by two uncorrelated factors”) can be tested. These methods need not be contrasted because one can explore the number of dimensions fitting data by comparing the outcomes of confirmatory analyses with plausibly alternative structures (e.g., one factor, two orthogonal factors, two correlated factors).

The approach to estimating dimensionality can proceed in either an exploratory or a confirmatory way (as described below). In situations where a concise summary of variables is the goal, and learning or estimating the dimensionality is not the goal, principal components analysis (described briefly below) is sufficient.

2.1.1 "Exploratory" vs. "Confirmatory" approaches

Typically, many latent variable modeling methods are classified as being motivated by either an exploratory approach, where the data and its inherent structure tend to drive the results and their interpretation, or a confirmatory approach, where a theoretically-motivated model is built and its fit to the data is estimated; the extent to which the model fits the data supports rejection or retention of the theoretical model. This conceptualization suggests that modeling must be classified as either exploratory or confirmatory, when in reality – and as implied by Thurstone (1937; 1940; 1947) - modeling more typically involves both exploration and confirmation (see, e.g., Dilalla, 2000; Hoyle, 2000; Nesselroade, 1994). Some argued that exploratory research requires theoretical motivation, and so cannot be "purely exploratory" (e.g., Harman, 1976; Spirtes, Glymour & Scheines, 2000), and others pointed out simply that most research is difficult to classify as purely exploratory or purely confirmatory (Bollen 1989; Dilalla 2000; Hoyle, 2000; Nesselroade, 1994).

Widaman (2007) articulated the differences between common factor analysis, which is an approach to estimating the off-diagonal correlations in the matrix of observed variables, and principal components analysis (PCA), a method for explaining as much of the variance (not covariance) among the observed variables as possible using fewer, composite, variables. These two goals were also identified by Harman (1976, p. 14) as the two distinct objectives of forming a linear combination of observed variables. Widaman (2007) noted that the common factor analysis model is "(t)he linear model that relates latent variables to manifest variables…"(pp. 185-6) –latent variables should already have been hypothesized by the investigator prior to the estimation of the linear model relating these to the indicators. Thus, the "exploration" is limited to discovering what observed variables are related to which latent ones, and to estimate the strength of those relations in the form of pathweights.

Conversely, the PCA approach seeks to condense the information so that the maximum amount of variability among the original variables is contained "…in an efficient, reduced-rank representation" (p. 187). In general, these are the two goals which current "exploratory factor analyses" are used to achieve, although they are essentially contradictory. Again, in common factor analysis and PCA, the "exploration" takes the form of estimating the optimal weights with which to combine the observed variables in the formation of new composites. In both cases, finding a zero weight, or that the contribution of an observed variable to the factor or composite is not significantly different from zero, signals a discovery that the observed variable is not related to the factor. Neither approach (common factor analysis, PCA) provides evidence of latent (i.e., not observed) variables that underlie the observed variables, especially as currently carried out with common software. Using factor analytic methods to determine (or rather, estimate) the number of latent factors that a set of observed variables represents will always fail in the sense that no model-fitness evidence can be obtained that none of the observed variables is actually responsible for the correlations of other observed variables (Heywood cases suggest an indicator is a cause, but this will not be reflected in a model fitness statistic). Although Velicer, et al. (2000) argue and demonstrate that two methods (minimum average partials, MAP (Velicer, 1976) and parallel analysis, PA (Horn, 1965)) are accurate for enumerating components (p. 68), components are not factors. Components are latent in the sense that they did not necessarily exist before the data were collected (since PCA estimates optimal weighting to create components from observed variables), rather than in the sense that they are reasonable causes for the observed correlations among variables.

2.1.2 Principal Components Analysis vs. Factor Analysis

It was clear when the two methods were first introduced that PCA (Spearman, 1904; Hotelling, 1933) was different from factor analysis (FA, Spearman (1904; 1927)), and this distinction has been maintained in some textbooks (Lawley & Maxwell, 1962; Cudeck, 2000; Stevens, 2002; Afifi, Clark, & May, 2004; Manly, 2005; and see Cowles, 2001, chapter 11), but confounded (Kline, 1994; Afifi, Clark and May, 2004, p. 395) or contradicted (e.g., Reyment & Jöreskog, 1996, p. 4; Johnson, & Wichern, 2002, p. 448; Pett, Lackey, & Sullivan, 2003, p.2) in others. It is possible that much of the confusion among researchers between PCA and FA arises from the variability across references in the description (or nomenclature) for these methods. Osborne, Costello and Kellow (2008) noted that software providing exploratory factor analysis uses default settings (e.g., using PCA to extract the factors and offering an eigenvalues-greater-than-one stopping rule) that are inconsistent with methods and decisions that are known to be "best practices" in this area (p. 87).

PCA and FA are distinct in terms of purpose, algebra, and interpretability. While the most popular conceptualization is that these are the same, they are not (see Widaman, 2007 for a recent and thorough differentiation), although these differences are typically blurred by confusing (Gorsuch, 1983: chapter 6; Reyment & Jöreskog, 1996; Kline, 1994) or possibly misleading (Johnson & Wichern, 2002, p. 448; Pett, Lackey, & Sullivan, 2003, p.2) representation in reference textbooks; other representations might rely on the readers' appreciation of the mathematical or algebraic distinctions shown (e.g., Bartholomew & Knott, 1999: chapter 1; Cureton & D'Agostino, 1983; Johnson & Wichern, 2002, chapter 9). In other cases, either the differences between the methods for factor analysis and principal components analysis are not highlighted, or methods are labeled in potentially confusing ways, such as "principal factors" (Gorsuch, 1983), "principal components factor analysis" (Manly, 2005, p. 95; Reyment & Jöreskog, 1996), or "orthogonal factors" (Johnson & Wichern, 2002: chapter 9). The differences between components (i.e., in the sense of Mulaik (1972), how the observed variables together represent an underlying one) and principal components (i.e., an extraction method) is also a source of confusion.

The results of PCA and FA can often be similar (Bartholomew & Knott, 1999; Stevens, 2002), PCA is sometimes used as a first pass to identifying the number of latent factors, and PCA is characterized and used as a method for extracting factors in an exploratory factor analysis (e.g., Velicer, et al. 2000, p. 68; although see Manly, 2005, p. 95) as the default approach in popular software packages (SPSS, SPSS Inc.).

These tendencies in modern uses of PCA in/as EFA are contrasted, but not very distinctly (nor very often), with the use of other methods for extracting a pre-specified number of factors (Bartholomew and Knott, 1999; Afifi, Clark, & May, 2004; Gorsuch, 1983; Harman, 1976). In what might be called "principal components analysis-factor analysis" (PCA-FA), principal components analysis is used to extract orthogonal factors, and the correlations between each observed variable and the factors (i.e., loadings) are estimated.

An alternative is principal axis factoring (PAF) which differs from PCA-FA in that the diagonal of the correlation matrix is replaced with communalities, rather than 1s, as estimates of each variable's variance, and also in that the number of factors can not be larger than one fewer than the total number of variables in the problem. In PAF the starting point of the extraction occurs after "unique" variance in the observed values variables has been removed from the diagonal values in the correlation matrix. Thus, PAF is a method for estimating the common components explaining only the covariance in a set of variables (see Mulaik, 1978 pp. 174-175). What is critical in PAF is that, while PCA-driven factor extraction follows the variance, that is, the components are highly sensitive to the variability in the individual variables being analyzed, PAF-driven factor extraction follows the covariance, that is, the components are sensitive to the levels of association among variables, and not to the variability of any one variable. This can lead to solutions that differ in critical ways depending on the extraction method – with PAF always tending towards explaining associations while PCA will tend towards explaining the variability in one variable at a time (particularly when there are discrepancies in the variabilities of the indicators).

In summary, the literature and specifically, reference texts, may be confusing to readers in terms of what common factor analysis actually is. Investigators who choose their factor analysis method on the basis of software, rather than their ultimate research goals, may consolidate their variables into a smaller space (i.e., PCA), or may estimate the correlations between observed variables and factors that best account for covariance (e.g., PAF), without (knowing whether they are) achieving their actual goals.

2.1.3 Estimating dimensionality using “stopping rules”

Whether the investigator chooses, or uses, PCA, PAF, or another common factor analysis method, the number of factors or dimensions can be chosen on the basis of many different “stopping rules”, or algorithms for determining when to stop extracting factors (Bandolos & Boehm, in press; Hoyle & Duval, 2004; Velicer, et al. 2000)). Each of these rules is inconsistent with a search for underlying causal structure, in the sense that they are all based on the explained variance, or on the similarities of the model-implied correlations to the observed correlations. That is, stopping rules are not based on evidence of the existence of latent causal factors. Stopping rules essentially summarize the fit of the data with one of the two purposes of the factor analytic enterprise (maximizing variance explained or minimizing differences between model-implied and observed correlations). Stopping rules do not provide direct information about the number of causal factors that appropriately characterize the measurement models. It could be argued that only a confirmatory analysis could do this (i.e., determining whether a model with more or fewer factors fits the data better), but maximum likelihood common factor analysis (Bartholomew

& Knott, 2000) solutions that improve in fit up to a certain number of extracted factors could also guide “stopping”.

PCA is characterized as a method for concisely explaining variance, while FA is characterized as a method for explaining covariance, among a set of observed variables (Cudeck, 2000, p. 275; Kenny, 1979, pp. 121-2). FA is not geared towards concision but rather, explanation (of the covariance by the factors), but as Mulaik (1972) cautioned, “Factor analysis is not a method for discovering full-blown structural theories about a domain.” p. xii).). This caveat highlights the inconsistency of using a stopping rule and a search for causal structure. In fact, there appears to be a distinct shift in the emphasis of earlier researchers in factor analysis and its methods (e.g., Cattell, 1978; Cureton & D’Agostino, 1983; Gorsuch, 1983; Harman, 1976; Mulaik, 1972; Thurstone, 1947). This emphasis on seeking support for theory using FA (or PCA) is missing from more current versions (e.g., Johnson and Wichern, 2002; Pett, Lackey, & Sullivan, 2003; Reyment & Jöreskog, 1996), which do not consider the origins or theoretical components that earlier factor analytic developers and researchers found to be so crucial to the explanation, use, and interpretability of factor analytic methods.

The issue of the correct number of factors was of concern to earlier investigators, but their approaches to this critical problem was more of an appeal to theory, and not an appeal to algorithms or fit of the factor solution (the model-implied covariance matrix) to the data (the observed covariance matrix). An example of the contrast in perspectives on this method is Johnson and Wichern (2002) statement that “Factor analysis can be considered an extension of principal components analysis” (p.

478), although Bartholomew and Knott (1999) pointed out that a key difference between these is that PCA involves no probabilistic structure (p. 13); neither of these presentations suggest a role for theory in the evaluation of a factor analytic solution. By stark contrast, Cattell (1978) noted, "…component analysis is no good as a final scientific model…" (p. 17), which is more in keeping with Thurstone's avocation that "…exploration with factor analysis required carefully chosen variables and that results from using the method were only provisional in suggesting ideas for further research." (paraphrased in Mulaik, 1978 (p. xii)).

The distinctions between the mathematics involved in PCA and FA are outlined by Cattell (1978), Gorsuch (1983) and Harman (1976) and are not explicated here. Neither method addresses the possibility that one of the observed variables is actually providing the common influence that both methods can detect. This makes PCA inappropriate, and FA less appropriate, approaches when the goal of the analysis is to understand, uncover or discover the latent structure present in a collection of observed variables (see Silva, et al. 2006). It should be noted that neither PCA nor FA is specific to the situation Wang, Wilson, and Adams (1997) refer to as between-variables multidimensionality, and what is often referred to as "simple structure", i.e., where each observed variable is associated with, or most influenced by, a single cause or factor. However, because simple structure is often a research goal, and because this *is* the only context in which vanishing tetrads (CCA) can operate, this dissertation is focused on estimating dimensionality when only one factor is sought for each observed variable. When this is the case, Kenny (1979) calls the observed variables "indicators"; other investigators use this label for any observed variable with a(t least one) latent cause. Because the feature of correlation constraint analysis that was pursued in this dissertation is focused on discovering “pure” factors, those where indicators have only one cause, observed variables are referred to as indicators. The next section presents the mathematics involved in FA and contrasts them with those involved in correlation constraint analysis (CCA).

2.2 Algebraic Representations of Common Factor and Correlation Constraint Analyses

Typical representations of the mathematical underpinnings of common factor analysis involve matrix-level manipulations (e.g., Gorsuch, 1983 Chapter 3; Skrondal & Rabe-Hesketh, 2004, chapter 3), showing the common factor model, which is described in section 2.2.1. The representation of correlation constraint analysis is typically given in terms of correlations, and this is presented in section 2.2.2.

2.2.1 The Common Factor Model

Yates (1987, pp. 9-15) gave the following derivation for the common factor model. The emphasis on *the* model comes from the fact that all representations of factor analysis, from different disciplines and perspectives, share the same starting point.

$$\mathbf{Z}' = \mathbf{P}\mathbf{F}' + \mathbf{U}\mathbf{E}'$$

where, with n indicators (n=1,…,j) representing N= 1, …i individuals and m common (latent) variables (m =1, …k),

$$\mathbf{Z}\ (\mathrm{N \times n}) = \{z_{ij}\}$$

is the matrix of observed standard scores for the N individuals (1, …, i) on the n indicators (1, …, j);

$$\mathbf{F}\ (\mathrm{N} \times m) = \{f_{ik}\}$$

is the unobserved matrix of standard scores for the N individuals (1, …, *i*) on the *m* (common) latent variables (1, …, *k*);

$$\mathbf{E}\ (\mathrm{N} \times n) = \{e_{ij}\}$$

is the unobserved matrix of orthogonal standard scores for the N individuals (1, …, *i*) on the *n* latent variables that are unique to the indicators (1, …, *j*);

$$\mathbf{P}\ (n \times m) = \{p_{jk}\}$$

is the matrix of loadings (regression coefficients) for the n indicator variables (1, …, *j*) on the *m* latent common factors (1, …, *k*); and

$$\mathbf{U}\ (n \times n) = \mathrm{Diag.}\ \{u_1, \ldots, u_n\}$$

is a diagonal matrix of the loadings (regression coefficients) for the *n* indicator variables, representing the variance in that variable that cannot be accounted for by (regression on) any common factor.

Yates continued (p. 11) by summarizing the common factor model in terms of the correlations among the observed variables (in an *n* x *n* matrix, **R**) in terms of what must be estimated in the common factor analytic method:

$$\mathbf{R} = \mathbf{PCP'} + \mathbf{U}^2$$

where **P** is the *n* x *m* matrix of loadings of the *n* observed variables on the *m* common latent variables outlined above, **U** is the *n* x *n* diagonal matrix of loadings of each observed variable on its unique latent variable (i.e., variance in the observed indicator that is unique to it) and $\mathbf{C} = \frac{1}{N}\mathbf{FF'}$ is an *m* x *m* matrix of correlations among the *m* common factors. The diagonal matrix $\mathbf{U}^2$ represents the contributions of each indicator's unique latent variable to the correlations among all the observed variables.

From this derivation, it can be seen that the estimation of **P**, the *n* x *m* matrix of loadings of the *n* observed variables on the *m* common latent variables, and by extension, of $\mathbf{U}^2$, the "uniquenesses" of the observed variables, is required for common factor analysis. It can also be seen that the dimensionality of the matrices must be specified for any solution – that is, the number of *m* common factors must be specified before estimation. Thus, as noted in Chapter 1, knowing the number of factors is crucial in estimating the associations of the indicators with those factors.

In more common, covariance form, the common factor model (given in Briggs & MacCallum, 2003, p. 26) is

$$\mathbf{\Sigma} = \mathbf{\Lambda}\, \mathbf{\Lambda}' + \mathbf{\Theta}^2$$

where **Σ** represents the (population) covariance matrix (corresponding to Yates' **R** matrix), **Λ** represents the loadings of observed on latent variables (in the population; corresponding to Yates' **P** matrix), and $\mathbf{\Theta}^2$ represents the matrix of uniquenesses (corresponding to Yates' $\mathbf{U}^2$ matrix). In ordinary least squares estimation of the **Λ** and $\mathbf{\Theta}^2$ matrices, the difference between the sample covariance matrix (**S**) and the covariance matrix implied by the common factor model ($\hat{\Sigma}$) is minimized according to the function

$$\mathbf{F}_{\mathrm{OLS}}(\mathbf{S}, \hat{\Sigma}) = \tfrac{1}{2}\mathbf{tr}[(\mathbf{S} - \hat{\Sigma})^2]$$

(Briggs and MacCallum, 2003, p. 28). The estimation procedure can also be described in regression terms, where the variability in an observed variable x_i, is predicted by its loading (λ_{i1}) on the common factor ξ_1, with uniqueness δ_i (Bollen & Ting, 1993 p. 150):

$$x_i = \lambda_{i1}\xi_1 + \delta_i .$$

This representation will help tie CFA to correlation constraint analysis, to be described in the following section (2.2.2).

An important feature of CFA that is not relevant in CCA is the factor loadings, and the role of loadings in the decisions about how many dimensions the data have (typically under simple structure). This was outlined in Chapter 1; where it was also noted that CCA does not yield estimates of the correlations between the latent variable(s) and indicators.

2.2.2 The correlation constraint model

A comparison of CFA and CCA in terms of their algebra is presented in Bollen and Ting (1993, pp. 150-153). Given a set of four observed variables (x_1, …, x_4) and a single latent variable ξ_1, and assuming that the average disturbance is zero (i.e., $E(\delta_i)=0$ for all four indicators and these disturbance terms are all mutually orthogonal (i.e., $COV(\delta_i, \delta_j)=0$ for any $i \neq j$) and are orthogonal to the latent variable (i.e., $COV(\xi_1, \delta_j)=0$ for all i), then the covariances among the observed variables in the population (σ_{ij}) would be computed as $\sigma_{ij} = \lambda_{i1}\lambda_{j1}\phi$, where ϕ is the variance of the latent variable ξ_1 (this assumes that the factor loadings were available). If the model generating these factor loadings and factor variance is correct, then the following equations must hold in the population:

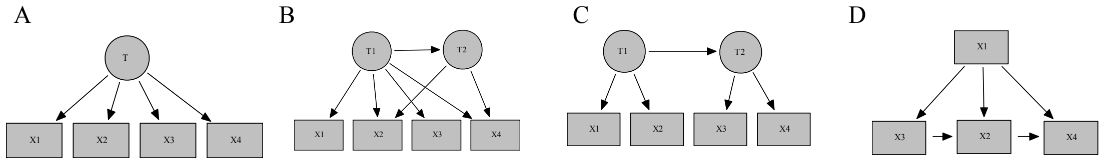
$$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} = 0$$

$$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} = 0$$

$$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} = 0$$

Spearman (1904; 1927) discovered these relationships; Kelley (1928) labeled these tetrad differences, or tetrads, $\tau_{ghij} = \sigma_{gh}\sigma_{ij} - \sigma_{gi}\sigma_{hj}$. When $\tau_{ghij} = 0$, the tetrad is called "vanishing" or is said to have vanished, and a vanishing tetrad means there is a common cause (of the observed covariances). Tetrads are computed as the determinants of all 2x2 covariance (sub)matrices (i.e., there must be at least four variables; in variance-covariance matrices larger than 2x2, determinants are computed for every set of four variables in the matrix) (Glymour, et al. 1987; Bollen, 1990; Bollen & Ting, 1993). A system of four observed variables –if they are associated (non-zero correlations)-will imply a set of three tetrad equations. The TETRAD software computes these tetrads for every set of four variables within the system being studied, to uncover evidence of common causes. TETRAD also computes partial correlations, with the influence of each observed variable partialled out, in order to determine if nonzero correlations vanish (change to zero) when an observed variable's influence is removed. Only vanishing tetrads with non-vanishing partial correlations are evidence of *latent* common causes. Figure 2 shows four models and the tetrad/partial correlation results that support conclusions regarding the structure of each model shown.

Figure 2. Tetrads and implied relationships between correlations. Circles: Latent causes; squares: observed variables.



A

$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} = 0$

$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} = 0$

$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} = 0$

& no partial correlations vanish.

B

$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} \neq 0$

$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} \neq 0$

$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} \neq 0$

& no partial correlations vanish.

C

$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} = 0$

$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} \neq 0$

$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} \neq 0$

& no partial correlations vanish.

D

$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} \neq 0$

$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} \neq 0$

$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} \neq 0$

& partial correlations involving X1 vanish.

Adapted from Scheines, et al. 1999; Figures 8 & 10, p. 183 & 185.

Since no sample values of tetrads can be expected to vanish exactly (i.e., the pairs of correlation products will not be exactly equal), Bollen (1990) developed a statistical test of whether, given a sample value for a tetrad difference, the difference is ‘significantly different from zero’ (a statistical test for determining whether partial correlations are statistically significantly different from zero is described in Glymour, et al. 1987, Ch. 11). Kenny (1979, pp 117-118) described an earlier test for vanishing tetrads involving canonical correlations; however, Bollen’s work on inferences for tetrad differences has been integrated into the TETRAD software and extended to confirmatory latent variable analysis (Bollen & Ting, 1993). The methods are robust and proofs have been published for the TETRAD software (Spirtes, et al. 2000) and for the tetrad differences more generally (Drton, et al. 2007). However it must be noted that this approach only works if there are at least four indicators, and that the TETRAD software has particular algorithmic constraints that may not be aligned with an investigator’s goals. The TETRAD program is described in detail in the documentation on the website (http://www.phil.cmu.edu/projects/tetrad_download), as well as in two books (Glymour, et al. 1987; Spirtes, et al. 2000).

When TETRAD finds that a tetrad *is not* statistically significantly different from zero, it concludes that a common cause exists. When TETRAD estimates that a tetrad *is* statistically significantly different from zero, it concludes that *no* common cause exists – but this can also occur when there are two or more common causes, since vanishing tetrads were developed specifically to detect single common factors (Glymour, et al. 1987). When it cannot uncover sufficient evidence of causality, it returns simply a representation of association (i.e., that two variables covary without specifying a common cause). TETRAD can be used to uncover measurement model families, and evidence of latent causal variables, but it is up to the analyst and content expert to build, fit and estimate (and then validate) reasonable models that could be based on the uncovered model families (Glymour, et al. 1987; Scheines, et al.1998; Silva, 2005; Silva, et al. 2006; Spirtes, et al. 2000). In this sense, CCA involves a dependence on theory that CFA (as done currently) might be argued not to exhibit.

2.2.3 Contrasting CFA and CCA

Consider panel A in Figure 2, a one-factor model with four indicators. The CFA representation of this model contains four equations of the form,

$$x_i = \lambda_{i1}\xi_1 + \delta_i$$

while the CCA representation of this model contains three equations of the form,

$$\sigma_{12} \times \sigma_{34} - \sigma_{13} \times \sigma_{24} = 0$$

$$\sigma_{13} \times \sigma_{42} - \sigma_{14} \times \sigma_{32} = 0$$

$$\sigma_{14} \times \sigma_{23} - \sigma_{12} \times \sigma_{43} = 0\,,$$

together with eight equations of the form**:**

$\rho_{12.3} \neq 0$ $\rho_{12.4} \neq 0$ $\rho_{13.2} \neq 0$ $\rho_{13.4} \neq 0$	$\rho_{14.2} \neq 0$ $\rho_{14.3} \neq 0$ $\rho_{24.3} \neq 0$ $\rho_{23.4} \neq 0$

Comparing these sets of equations, and given the forgoing, it can be seen that CFA involves the estimation of correlations between observed and unobserved variables, which can be achieved in OLS estimation as a function of having articulated the ‘correct’ number of unobserved variables. By contrast, CCA requires the estimation of correlations and partial correlations among the observed variables only.

TETRAD searches for evidence of causal structure using algebraic relationships, or constraints, that are implied by causal structures. These polynomials are termed “invariants” in the mathematical/algebraic literature (Drton, et al. 2007) and their existence, as a representation of a common cause, was discovered and described by Charles Spearman (Spearman, 1904; 1927); Kelley (1928) described pentads which, like tetrads, can identify common causes; while tetrads only identify a single common cause, pentads identify two common causes (pentads are described in Drton, et al. 2007)).

Critically, if more than one latent factor causes observed correlations, vanishing tetrads cannot detect the second factor. Applying the same logic to vanishing pentads (as exists in the TETRAD software) would identify the two latent causes, but to date no pentad-based analytic software exists. Another critical point about the TETRAD software is that, although there are modules that accommodate latent variables that share indicators with other factors ("impure" factors), the algorithm this dissertation utilized, BuildPureClusters does not accommodate impure factors. That is, each indicator can be assigned to only one latent variable, when sufficient evidence of the assignment exists. Furthermore, the algorithm by which the TETRAD software searches the covariance structure is not fixed; repeating the analysis on the same data is not guaranteed to find the exact same structure. Thus, CCA by TETRAD is, in many different senses, dependent on the TETRAD algorithms.

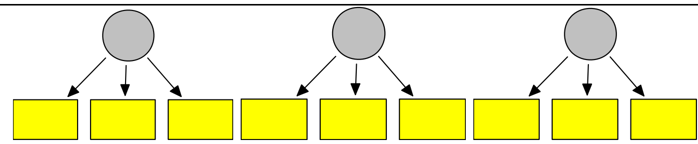
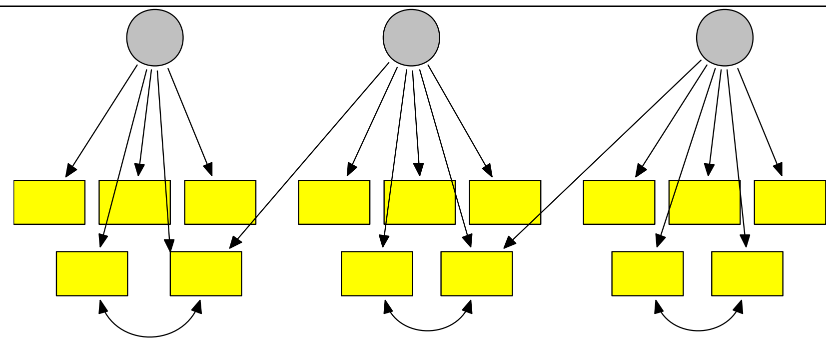
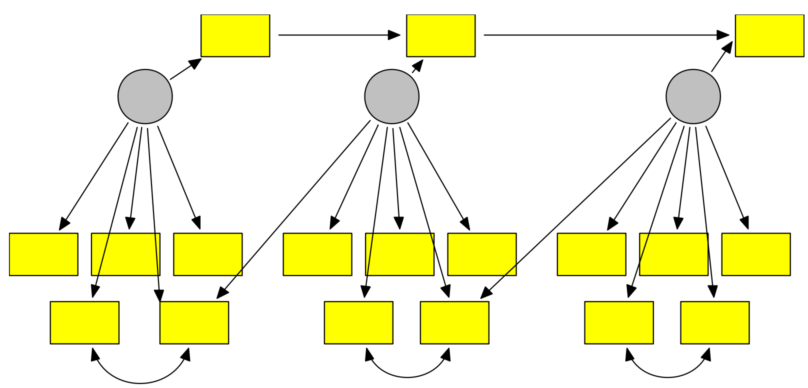
The tetrad equations represent causal pathways and importantly, the sample space is searched by TETRAD in a wholly unguided manner, that is, by algorithm. Thus, it might be considered that TETRAD proceeds from a position of ignorance in a similar sense that CFA does: TETRAD uncovers evidence of latent variables by algorithmic (not theory-driven) search of the sample space (theory-driven) while uncovering dimensionality by CFA relies on estimating correlations (not theory-driven) between observed (theory-driven) and unobserved variables. Both methods begin with information obtained with theoretical considerations, but proceed in data-driven ways.

2.3 Previous Simulation Studies Comparing FA and CCA

Previous simulation studies comparing CCA and exploratory factor analysis Silva (2005; Silva, et al. 2006) described a simulation study directly comparing the new TETRAD search algorithm (which his 2005 doctoral thesis established, and which the 2006 manuscript outlined), BuildPureClusters, against the exploratory factor analysis function of R, *factanal*. *Factanal* carries out ML estimation – common factor analysis (R Help, search on "factanal"), making the Silva simulations most closely representative of the research described in this dissertation. Silva's simulation involved 10 trials with sample sizes of 200, 1000, and 10,000. The purpose of Silva's simulation work was to uncover both the measurement model and the structural model; since only measurement models are considered in this dissertation work (i.e., dimensions, and not how the dimensions are related which would be the structural models), these are the features of the Silva simulations discussed here. Table 1 shows the three measurement models that drove the simulations considered by Silva.

Errors were classified as latent omission (number of latent variables not identified in solution (i.e., missing according to true model) divided by true number of latents); latent commission (number of latent variables identified in solution that were never in the true model, divided by true number of latents); misclustered indicators (number of indicators that were assigned to the wrong latent divided by the number of indicators in the true model); indicator omission (number of observed variables not assigned to any latent variable (missing according to true model) divided by number of indicators in the true model); and indicator commission (number of indicators observed in factors where they were not divided by number of indicators that did belong). Identifying "which factor" a set of indicators represented was based on the greatest sum of absolute valued-loadings; if all indicators loaded on one factor, that factor would be "identified" as the one with the highest sum of loadings and all other indicators would be classified as indicator commissions.

Table 1. Three measurement models used by Silva (2005) to compare CCA and ML-CFA.

	Three 3-indicator latent variables.
	Three 5-indicator latent variables with cross-loadings and correlated errors.
	Three 6-indicator latent variables with single causally-associated indicators, plus cross-loadings and correlated errors.

From each of the three measurement models (combined with three structural models that are not described here) shown in Table 1, ten samples were generated at each of the three sample sizes (200, 1000, 10,000). In every sample, pathweights were not fixed but were randomly generated from particular intervals –that is, every sample (irrespective of underlying model) had a randomly specified set of relations between indicators and their associated latent variables.

However, Silva (2005) also did not count any cases where fewer than four indicators loaded on any factor or were otherwise inconsistent with the other TETRAD modules or algorithms (i.e., whether there were omission or commission errors for factors or indicators, if they occurred in an improper solution, they were not counted).

Scoring of measurement model discovery performance by TETRAD (version IV) and *factanal* (R) were computed as the average (over ten samples). Tables 3.3 and 3.4 in Silva (2005, pp. 60-61) reflect the performances as a function of the combined measurement and structural modeling. Overall, TETRAD (BuildPureClusters) committed the smallest number of errors (on the order of about 5% or smaller, particularly for N=1,000 and N=10,000); the performance of CCA was found not to vary much across any of these design features. CFA was found to exhibit “very high” rates of latent commission (finding too many latent variables, between 2% and 97% of the time depending on the structural complexity and sample size). CFA tended to omit latent variables (find too few dimensions) on the order of 2% of the time. Of interest, a third CFA procedure was included in Silva’s simulation: a “purified” CFA result. In this approach, each indicator was only allowed to load on a single latent variable (irrespective of the true model). This might be comparable to a ‘simple structure’ approach to uncovering dimensionality; this method led to omission of latent variables (too few dimensions) on the order of 10%-60% of the time; and to commission (too many dimensions) 3%-20% of the time.

This dissertation extended Silva’s work by adding the element of consistency in the comparison (rather than using an average error rate), and also by examining the effects of strengths of associations, the presence and absence of relatively weak factors, and the inclusion of sample sizes that are more consistent with sizes applied researchers might utilize. Furthermore, by comparing CCA to CFA using OLS estimation, this dissertation involved a CFA method more similar to CCA than previous simulations (which all used ML estimation, shown by Briggs and MacCallum (2003) to be more error-prone than OLS in estimating weaker path loadings when stronger path loadings are present).

While not a direct comparison of CCA and EFA/CFA, Spirtes, et al. (2000) report a simulation study comparing TETRAD (version II) to EQS and LISREL in the context where a structural equation model (SEM) has been fit, and further paths are sought (Spirtes, et al., 2000, Ch. 11). In this simulation there were nine different 'true' SEMs simulated, and three incomplete versions (one for each of two of the true models and one for seven of the true models) from which TETRAD, EQS and LISREL started their respective searches for missing paths. From each of the nine true models, 20 samples with N=200 and 20 with N=2000 observations were generated. When sample sizes were 200, TETRAD identified the correct (complete) model, given the starting model, 52.2% of the time, while EQS and LISREL were correct 10.0% and 15.0% of the time, respectively. With N=2000, TETRAD provided the correct (complete) model 95% of the time while EQS and LISREL were correct 13.3% and 18.8% of the time, respectively.

2.3.1 Including a condition with zero factors

None of these simulations, and typically EFA simulations in general, included a case or condition with zero factors (uncorrelated observed variables). This is an important note, since EFA will extract a single factor, regardless of the extraction method. As can be seen in the derivation of the CFA estimation in Section 2.2.1, if m=0, no estimation can take place. Thus, even if there is no common factor underlying the observed variables, EFA must “find” at least one. By contrast, TETRAD can yield a solution with zero factors in two different ways: it can find evidence that no latent common causes exist, or it can find insufficient evidence that one latent common cause exists. These are three important possible outcomes of an analysis on a set of uncorrelated observed variables (one factor, evidence of no factors, insufficient evidence of factors). This dissertation sought to examine the two estimation approaches in the zero-factor condition, even though the two zero factor results from CCA could not be differentiated.

2.4 A note on establishing dimensionality

Without appeal to ‘minimization’ or concision in the resulting system, Pearl (2000), Spirtes, Glymour, and Scheines (2000) and Bartholomew and Knott (1999) articulate that identifying the latent variable that leads to local independence (independence of observed variables conditioning on the latent variable that is their common cause) leads to the discovery of a dimension or common cause (also called a common factor); this situation (conditional independence in a measurement model/causal model framework) is known as “d-separation” (Verma & Pearl, 1988)- but d-separation (the ‘d’ comes from the direction, or causality, in the model or graph under consideration) is specifically articulated for systems of observed variables only. Seeking the minimum number of dimensions required to achieve d-separation (as defined/described by Pearl, 2000) is neither the same method nor motivation as seeking the smallest observed (empirically supported) number of dimensions or common factors. These may also be contrasted with seeking evidence of latent common factors, since there is no impetus to find the smallest number, or "meaningful" factors.

The tetrads that are equal to zero suggest different patterns of association that explain the observed correlation matrix. Critically, tetrads will vanish (are zero) when the relationships are present without depending on the size of the correlation (Drton, et al. 2007). This is a crucial difference between using tetrads (see Spirtes, et al. 2000) to uncover latent structure and using other methods, such as CFA: the estimation in CCA is restricted to sample values of correlations and partial correlations. In CFA, the estimation is at the level of the correlations between observed and unobserved variables. The number must have been chosen *a priori* –see Section 2.2.1), and the loadings are estimated jointly while in CCA (using TETRAD) the estimates proceed four variables at a time.

However, the sample correlations will vary more when the sample sizes are smaller, and TETRAD algorithms appear to be most effective when population correlation values are at least 0.30 (Spirtes, et al., 2000). Small samples (N=100) and small correlations (0.25) were included in this study to compare the CCA and CFA performance in those conditions. These different types of estimation are highlighted here only because they constitute the crux of differences between the two methods, and not as an argument that one method is superior to another. Further, the data simulated in this study were characterized to challenge the two methods; no effort was made to map out the data characteristics for which one method was superior, or to determine at which point(s) either method would no longer be sensitive to any of the characteristics. As a first direct comparison of the two methods in realistic data contexts, the objective was to compare the error types and rates as described in Chapter 3.

In summary, uncovering the causal structure, or “true dimensionality”, of a set of variables, has been a motivating force in the development of factor analytic techniques for over 100 years. In 1987, Glymour, et al. developed a set of algorithms, and software to implement them (which has been revised and augmented in the ensuing 20 years) that automates the identification of vanishing tetrads and non-vanishing partial correlations, which can be used to uncover evidence of causal structure. This automated search software is not well known to many researchers, nor has its performance, relative to OLS estimation, been examined along the dimensions outlined in Chapter 1 and explicated in Chapter 3. Since vanishing tetrads imply common causes (causes that d-separate other variables), and latent variables (causes) are only identified when observed variables are ruled out as being the common cause of variability in other observed variables, it was important to determine whether this software performs as well as CFA by OLS, or better, because if so, it could bring a new perspective to uncovering dimensionality when indicators (mainly) have single causes, such as was the case in this study. The next Chapter describes how the research was carried out to test and compare the performances of these methods.

Chapter 3: Dissertation Methods

Chapter 1 outlined the research question driving this study: is CCA as good as CFA at uncovering dimensionality? It was specified that the context for testing this question was the simple structure, or between-variables multidimensionality, case, where each observed variable is influenced by only one latent variable, although impure factors represents one of the challenges this study put to the methods. This chapter outlines the methods by which this question was answered. Specifically, whether CCA and CFA performed at similar levels of accuracy (obtain the correct number of latent variables, correctly assign observed variables to their corresponding latents) and the level at which the analytic solutions were replicated (consistent) over repeated trials. These two performance characteristics (accuracy, consistency) were examined in contexts where sample size, model complexity, the presence of weak common factors, and correlated vs. independent latent variables were varied, as well as in a zero factor condition.

3.1 Design features

Models representing N=100, N=300, and N=500 observations were simulated that varied in the number of factors (0, 1, 4, 6), strengths of associations between indicators and factors (high (.80), medium (.60) and low (.40), see note in "scoring", below), saturation of latent variables in the model (pure, share one, and share two indicators), the presence or absence of weak common factors (weak=factor loadings are 50% of the strength of the other factor loadings (after Briggs & MacCallum, 2003), and structure (factors independent, factors moderately correlated (0.50

“moderately” correlated factors (French & Finch, 2006))). To focus the simulation, all data was generated from multivariate normal distributions, the models were all assumed to hold exactly in the population, and all factors had four indicators. Four indicators per latent is the minimum for TETRAD and will ensure that all measurement models are identifiable (Kline, 2005) but do not yield just-identified measurement models in cases where two indicators are shared between two factors. Four indicators is a realistic number for some measurement models (particularly for shorter scales, or sets of scores) that should be challenging for both of the two dimension-estimation methods, particularly when one and two indicators are shared by two factors.

3.2.1 Data Simulation

A set of R programs (http://incanter.org/src/R/modeler.R; http://incanter.org/src/R/modeler.test.R; http://incanter.org/src/R/cfa.joreskog69.R; http://incanter.org/src/R/cfa.joreskog69.test.R; D. Liebke, personal communication September 2008) were written for model-implied covariance matrices (and their analysis) in Jöreskog (1969). This program accepts the user’s latent variable model (graph) with specified paths (loadings). From this input, and using the Jöreskog (1969) equations and approach, a model-implied covariance matrix is generated, which becomes the population covariance matrix for the simulation. The random normal variable generator function of R (*rnorm*) generates the number of standard normal variables the user specified as indicators in the graph as input; a Cholesky decomposition of the population covariance matrix (according to the user’s model) is multiplied by the *rxc* matrix of *rnorm*-generated variables (where *r* represents the sample size and *c* represents the number of indicators). Graphs (Pearl, 2000) were created representing the conditions outlined above, and 50 samples were obtained from each of these graphs at each of the three sample sizes (N=100, 300, 500) described below. These represented fifty trials for TETRAD and for the common factor analysis using CEFA (Browne, Cudeck, Tatenini, & Mels, 2008). Fifty replications of each set of conditions were used to derive a performance summary for each method at each simulation configuration (outlined below). The result from each set of 50 trials was a mean number of errors (see Scoring, below) and the standard deviation for this mean. Fifty trials are more than any other CCA simulation has used, and is consistent with other simulation studies (e.g., MacCallum, et al. 2001; Preacher & MacCallum (2002)). Because a sampling distribution, or consistency with asymptotic assumptions, was not sought in this study, a larger number of replications was not warranted.

The same data were analyzed by the two methods.

3.2.2 Sample size

The three sample sizes (100, 300, 500) were chosen because they represent the levels of sampling variability used by Briggs and MacCallum (2003); they also can be considered to roughly correspond to 'small', 'medium' and 'large' samples. Silva (2005) used samples of size 200, 1,000 and 10,000, but the results for CCA and CFA did not vary between 1,000 and 10,000 and 1,000 is not a particularly realistic sample size in social science research. MacCallum, et al. (2001) used 60, 100, 200 and 400. In the present context, the question of interest is whether a 'larger' or a 'smaller' sample size will affect recovery of dimensions, so strictly speaking, the exact choice of sample sizes is not important. Sample size did affect performance in Silva's comparison of CFA (by ML) and CCA (p. 60-61), where he had manipulated only model complexity (and had not used OLS estimation) and found that CFA performed worse at N=200 than N=1,000, but the performance of CCA did not vary with sample size. Silva (2005) and Silva, et al. (2006) only reported the average error rate over 10 replications, and did not include features of accuracy and invariance as is proposed in this dissertation. Briggs and MacCallum (2003) found that OLS performed well whether the sample size was 100 or 500, but ML performed less well than OLS when the sample size was 100. This may be material in the present context since CCA has only ever been compared to ML (at whatever sample size); it is unclear if an advantage that CCA exhibits over ML at smaller sample sizes (e.g.., 200 in Silva's 2005 study) will be lost when CCA is compared to CFA by OLS.

The choices of data characteristics were made specifically to challenge these methods and lead to error rates, rather than success rates, to be compared.

3.3 Simulation Design

To summarize this simulation design, this was a 4x3x3x3x2x2 factorial design (#factors x sample size x strength of association x saturation x weak factors x structure). The number of factors being 0 (observed correlations are a function of sampling variation) or 1 did not permit saturation, weak factor, and structural conditions; thus, there are 3x3x3x2x2x2 (=216) + 3 (zero factors at 3 sample sizes) + 9 (1 factor with 3 strengths and 3 sample sizes) = 228 different models were simulated. Graphical representations of most of these features appear in Figure 3 below.

Each model yielded 50 simulated samples for the N=100, N=300 and N=500 sample size conditions. These simulation factors all constitute functions of performance explored **within** each of the dimension-estimation methods.

3.3.1 Outcomes of Interest: Within Method

Accuracy of CFA (OLS estimation) and CCA was estimated based on three features: the primary outcome of interest was percent (of the 50 trials in each of the three sample size conditions) where the correct number of latent factors was identified – this represents the recovery of the correct dimensionality of the model. These percentages were averaged across conditions within method. Two additional accuracy outcomes are the percent where indicators are correctly identified as 'belonging' to their latent parent, and the percent where both of these features are correct.

3.3.2 What Constituted an Error

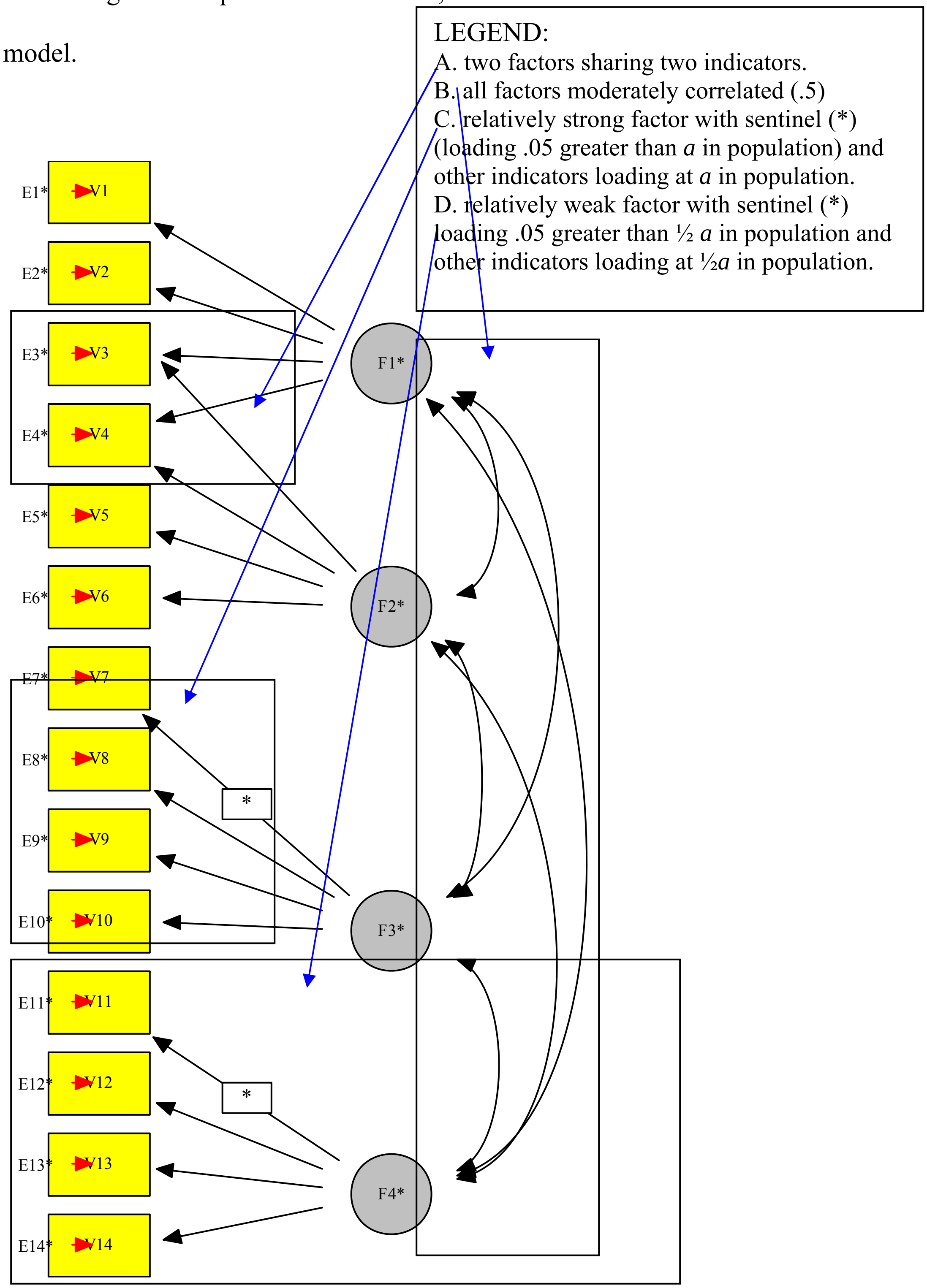
Performance of each method on each of 50 runs at the 228 different model configurations was quantified in terms of the percent (of 50 trials) where the correct number of latents was recovered, the percent where all indicators were correctly attributed to their true latent parent, and the percent where both of these were correct. The means and standard deviations for the 50 trials were recorded as the result for that condition. Errors were characterized as "commission" if they put an extra in (path or latent variable), errors were characterized as "omission" if they left out (path or latent variable).

All latents (L1, L2, …, L6) had four indicators (X1, X2, …, X4, X5, …, X16); however in the ‘share’ conditions, one or two indicators loaded with equal strength on two of the stronger factors, leading to one or two fewer indicators than a multiple of four) as described. Of these four indicators, the first one was the 'sentinel' for that factor: it had a slightly higher loading in the population, while the other three indicators had a slightly lower loading in the population. For example, in the stronger loading condition (referred to above as the .8 condition), the sentinel had a population loading of .85 while the other indicators had a population loading of .75. In the middling condition (referred to as the .6 condition), the sentinel loaded with .65 and the nonsentinels loaded at .55 in the population. In the weaker loading condition (referred to as the ‘.4’ condition) the sentinel loaded with .45 and nonsentinels loaded with .35 in the population. In all cases, as articulated above, the single weaker factor, if present, had population loadings of ½ the size of the stronger factors. The weak factor had four sentinels and nonsentinels defined the same as above, but with values 50% of those on the non-weak factors. No weak factor was involved in sharing (see Figure 3).

In cases where 2 or more sentinels load on a single factor, if there are other indicators (non-sentinels) associated with one of the sentinels, the majority of indicators will determine “which” factor it is, the other indicators will be treated as errors (of commission). If 2 or more sentinels load on a single factor and there are no non-sentinel indicators, the first sentinel was (arbitrarily) ruled to represent the latent factor and the other sentinels were ruled errors of commission.

Errors of omission are indicators that are left out of the solution (one error point). However, an error of commission can yield *two* points: if the indicator was assigned to the incorrect factor and no other factor, this represents both an error of commission (indicator on wrong factor) and an error of omission (indicator missing from correct factor). For example, X1 is the sentinel for L1 and X5 is the sentinel for L2. If X5 appeared on L1 (and on no other factors), it represents a one-point error of commission of X5 on L1 plus a one-point error of omission on L2, where X5 belonged. This is the only type of error of commission within the TETRAD procedure. However, in CFA, indicators could be found to load (defined and quantified as [estimated loading/SE]>1.96) on more than one factor. Thus, while errors of commission could only be worth two points in a solution from TETRAD, if CEFA assigned X5 to load on both L1 (commission) and L2 (no error), then it would only be a one-point error. With this scoring algorithm, TETRAD could incur 1-point errors of omission and could only incur 2-point errors of commission, while CFA could incur 1- and 2-point errors of commission and the 1-point error of omission that both methods could incur. Previous CCA studies have not resulted in dramatic error rates, so it was not thought likely that exploring the error types would be worthwhile. However, as was seen in the results, the reason why error rates were so low in other studies was the disallowed solutions that were improper – in the present study, error rates were obtained from every solution, rather from only those that were proper (for TETRAD).

Figure 3. Depiction of saturation, weak factors and structure for a 4-factor model.



Error rate scoring is described fully in Appendix 2. A series of one-way ANOVAs was used to determine sensitivity to these design features in terms of differences in the error rates (factor and indicator omission and commission errors) within-method and across the design features individually as reflected by effect sizes (described below). No effort was made to correct for multiple comparisons since, with large enough samples, it is likely that any effect could have been made significant; instead the within-method ANOVA results were treated as characteristics of each method, which could then be contrasted across methods.

Invariance in latent variable results (discovery of dimensionality) was operationalized as a within-method outcome as (yes/no) observing 95% of the same solution (number of latent variables, whether accurate or not) within each set of 50 analyses. A within-method failure to replicate accuracy and invariance results across all sample size conditions was taken to represent a sensitivity to sample size for accuracy and invariance of dimensional discovery by that method.

3.3.2 Outcomes of Interest: Between Method

The features were compared across methods, adding an additional factor for the overall study design. There were a total of four outcome variables to be extracted from each 100-sample 'trial' under each method, and invariance were operationalized across methods as the proportion of solutions that were the same out of the 11,400 trials at each combination of simulation features. For each of the four continuous outcomes, multi-way factorial ANOVAs were carried out, comparing CFA and CCA on all relevant simulation design features (main effects) and their interactions. The ANOVA results were summarized using effect sizes (η^2), representing an estimation of the relative proportions of variance that each of the simulation features explains (Sheskin, 2004). As Sheskin points out (pp. 913-914), there are two methods for computing η^2; one estimates the proportion of variability attributable to the factor relative to all other factors and their interactions, as well as within-group variability, while the other ("partial η^2") estimates the proportion of variability attributable to the factor relative to the within groups variability only, collapsing over the contributions from other factors and their interactions.

Table 2 summarizes the research question(s) of interest in terms of comparing performance (on the four performance outcomes) by multi-way factorial ANOVAs.

Table 2. Simulation features/ANOVA results to be estimated (for each outcome) based on simulation features.

Design characteristic (values)	Question of substantive interest
Main effects	
Method (CFA, CCA) (M)	Is the outcome (accuracy, invariance) significantly different depending on method? Collapsing across the other main effects, there are 228 observations per method in this cell (one outcome per 50 trials in each model configuration)
Sample size (100, 300, 500) (N)	**
Number of factors (0, 1, 4, 6) (F) *	**
Strength of relations (.8, .6, .4) (L) <see note on sentinels>	**
Saturation (pure, share 1, share 2) (S)	**
Weak factors (yes, no) (W)	**
Structural model (independent, correlated factors) (I)	**
Interactions (only interactions involving the method are of interest- no others will be tested)	

M x N	Do differences in method on outcome vary by sample size? (72 observations per method in this cell)
M x F	Do differences in method on outcome vary with # of factors?
M x L	Do differences in method on outcome vary depending on loading strength?
M x S	Do differences in method on outcome vary when indicators are pure or load on >1 factor?
M x W	Do differences in method on outcome vary in the presence of weak factors?
M x I	Do differences in method on outcome vary when factors are independent vs. correlated?
Multiway interactions: never reported (to date) for CCA vs. CFA, but MacCallum, et al. (2001) found no significant effects (based on η^2). Interpretation could be complex and the cell sizes will be small, but these might suggest design features worth pursuing in future research (focused on specific interactions, and their effect sizes).	
M x N x F	
M x N x L	
M x N x S	
M x N x W	
M x N x I	
M x N x F x L	
M x N x L x S	
M x N x S x W	
M x N x W x I	
M x N x F x L x S	
M x N x L x S x W	
M x N x S x W x I	
M x N x F x L x S x W	
M x N x F x L x S x I	
M x N x F x L x S x W x I	
* all factors have four indicators. ** These main effects are only of interest for their potential contributions to interactions with method.	

3.4 Estimation Features

CCA and CFA are very different methods (as articulated in Chapter 2), and one way to minimize this difference is to choose the estimation procedure for CFA that is more similar to how CCA works. As noted above, choosing CFA by a method like OLS, as opposed to ML (which was used in all other comparisons of CCA and standard CFA (Silva, 2005; Silva, et al. 2006) or search (model respecification; Spirtes, et al. 2000, chapter 11) simulations), would represent a fairer comparison of the performance of CCA and CFA, since the tetrad inequalities (outlined below) "… hold regardless of the values of the path coefficients and the variance of the latent variable." (Bollen & Ting, 2000, p. 7). This is not of particular concern for the zero and one factor samples, since Briggs and MacCallum (2003) reported similar performance by OLS and ML when no weak factor was present, and Akaike Information Criterion (AIC, Akaike, 1973; Konishi & Kitagawa, 2008) values can be computed using the log likelihood function or model χ^2.

Based on the presentation of estimation in Chapter 2, it can be seen that CFA models cannot be estimated with m=0 (although this *could* be approximated by estimating a multivariate normal distribution with a diagonal covariance matrix). To maintain a factor analytic focus in the non-TETRAD analyses, confirmatory factor analysis fitting a one factor model was carried out by Mplus v. 5.2 (Muthén & Muthén, 2008) using maximum likelihood estimation in order to obtain AIC estimates for all the zero and one factor simulated data (zero factors, four indicators, N=100, 300, 500 and one factor, four indicators, N=100, 300, 500 at each of the loading strengths (weak, medium and strong)). Standard errors are not available from Mplus if unweighted least squares (ULS, their version of OLS) estimation is used and these were required for assigning indicators to factors (loading/SE).

An interesting possibility with confirmatory analysis is that a one factor model might fit better (by AIC) than the null model while none of the loadings is significant (i.e., all loadings p>0.05). Whenever a factor analysis (confirmatory for 0 and 1 factor data, or common for 4 and 6 factor data) result found a factor, but no loadings significantly greater than zero on that factor, that factor was treated as "found" while indicators were treated as omitted (four errors of omission for the indicators, for each of the 50 trials).

Ordinary least squares estimation of factor loadings in CFA by CEFA was used for two reasons: a) Briggs and MacCallum (2003) found that OLS performed better than ML when weak factors were present; and b) Briggs and MacCallum (2003) noted that OLS does not differentially weight residuals depending on the strength of the correlations in the model, and this is similar to what CCA does, since the functions of the 2x2 submatrix determinants are proportional, rather than absolute. Although OLS is not strictly an appropriate estimation approach for latent variables (Bollen, 1996; Kline, 2005), it is implemented in CEFA (Browne, et al. 2008). CEFA v. 3.02 (Browne, et al. 2008) was used, with a SAS v.9.2 (SAS Institute, 2008) macro to run the CFA analyses. OLS estimation was used and geomin rotation was employed to obtain simple structure (i.e., so that indicators would tend to load on only one factor). Geomin, an oblique rotation, is one of those recommended as a "best practice" in exploratory factor analysis (Osborne, et al. 2008) and was also recently found by Asparouhov and Muthén (in press) to provide the best solutions in their large-scale simulation study. Oblique rotations were recommended (by investigators from Thurstone (1947) through Jennrich (2007)) as most realistic; when factors are orthogonal, the oblique rotation finds the right structure and when factors are oblique (as in half the simulations in this study), orthogonal rotation cannot find the right structure.

Akaike Information Criterion (AIC, Akaike 1973; Konishi & Kitagawa, 2008) values were computed for each of the CFA solutions [although the AIC computations are based on maximum likelihood and the likelihood is currently typically computed using ML estimation, any estimator could be utilized in the likelihood, since the goal is to replace Θ with $\hat{\Theta}$; the OLS estimator may be less efficient than the ML estimator for theta, but our formula for AIC can be computed with either $\hat{\Theta}$]. Anderson (2007: 60) recommended using a formulation of AIC with a second-order correction for bias (AICc, after McQuarrie & Tsai, 1988), but the AICc correction (relative to AIC) is a 'small sample' correction, while all the models under consideration within each CFA run will have the same sample size and varying complexity. The model selector (i.e., AIC) needed to be consistent for all models and the sample sizes of 100 were not “very small”, so AIC was to be preferred over AICc. Anderson (2007: 160-1) also noted that the Bayesian information criterion (BIC; Schwarz 1978) does not assume that the true model is in the set under consideration; and that in the face of increasing model complexity, unless the sample size is also increasing, "BIC does not guarantee a good parsimonious model..." The choice of AIC over BIC was made to promote consistency for model choice within a condition where only complexity was varying, since that is how the information criterion was utilized.

Consistent AIC (CAIC, Bozdogan, 1987) was compared to AIC by Anderson, Burnham & White (1998) and found to (incorrectly) select smaller models than AIC. Because Silva (2005) found that factor analysis tended to lead to larger models than TETRAD (BPC) did, CAIC was not chosen for model selection by CFA because it might offset the 'natural' tendency of CFA to over extract (relative to TETRAD). Conditional AIC (cAIC, see Wager, Vaida & Kauermann, 2007) is used for mixed (effects) models, and so is not appropriate for the current context. Thus, AIC was chosen as the criterion for identifying the “best” CFA result.

For CFA analyses on four and six factor number conditions, the solution with the best AIC was deemed “the winner” and that model’s solution (# latent variables, which indicators load on which latent variables, and the combination of these) was obtained as its performance characteristics.

In order for CFA results to match CCA results most closely, any pathweight that was statistically significantly different from zero was considered to represent that indicator “loading” on that factor. CCA algorithms can be adjusted so that the p-value for any test (of vanishing tetrads) accommodates the multiple comparisons (all possible tetrads, excluding those that are implied –and so not tested, Spirtes, et al. 2000; p. 271), but CFA pathweight tests are NOT corrected for the simultaneous estimation of loadings and their associated standard errors (see, e.g., Mplus 5.2 Users Guide, 2008). All tests of significance (i.e., were tetrad differences significantly different from zero in TETRAD and were loadings significantly different from zero in CEFA) were done at a testwise α=0.05.

The CFAs for the 4 and 6 factor samples were run extracting from 1 to the correct number+2 of factors (i.e., 1-6 factors for the four factor condition and 1-8 factors for the six factor condition). This represented a fairer comparison with TETRAD, which was observed to return as many (but not more than) two factors more than there were in the population. Over 68,000 CFA models were fit; the model

χ^2, number of parameters and degrees of freedom were obtained from each fit (using a combination of SAS and Excel macros). Determination of which indicators were assigned by CEFA to which factor, for each of the 50 replications within each condition, was taken from the one CEFA-generated result with the best AIC (computed as $(\chi^2/n\text{-}1) + (2k/(n-1))$, where χ^2 is model χ^2, n is the number of subjects, and k is $(.5v(v+1))\text{-}df$, where v is the number of variables and df is degrees of freedom). Maruyama (1997, pp. 246-7) noted that there are multiple formulae for computing AIC values and while they do not always yield the same value, they do agree in terms of ranking models. This formula was selected because it will provide the same ranking as the to provide the same ranking as the Expected Cross-Validation Index (ECVI, Browne & Cudeck, 1993) that is computed by CEFA for each model. This way, the results from Mplus (with AIC values computed with this formula) and the results from CEFA could be most directly comparable.

For the CFA model (based on the same one of the 50 replicates) with the best AIC, all loadings were transformed so that (loading/SE>1.96) = 1 and (loading/SE<1.96)=0, where 1= indicator loads on the factor and 0= indicator does not load on the factor.

Factor loading values themselves were not used to determine whether indicators were or were not assigned to any factor. That is, assignment of indicators to factors by CFA was reduced to the same present/absent characterization that TETRAD provided. In the study describing the BuildPureClusters algorithm (Silva, 2005), the sum of absolute values of loadings of those indicators truly associated with a given latent variable were used to determine which true latent variable an assigned mix of indicators represented. Apart from this, and the use of these absolute value sums to determine which true latent a found factor corresponded to most closely, factor loadings were not used by Silva (2005). The standard errors of factor loadings were not obtained or used by Silva (2005), nor were model fit (AIC) values.

CCA in TETRAD uses simple sample-estimated covariance matrices (and their associated 2x2 determinants) to compute the tetrads (and determine whether they are statistically significantly different from zero, as described in Chapter 2. Similarly, the partial correlations test (whether non-zero correlations are statistically significantly different from zero when an observed variable is partialled out) is based on sample covariances/correlations. The Wishart test for vanishing tetrads (Bollen, 1990) was used as implemented in TETRAD version IV.

3.4.2 Data Collection

All indicators were labeled “X#” and all latents were labeled “L#”. For example, the 3rd of four indicators loading on the second of four latent variables was X3 on L2. From each of the fifty trials within each of the 228 conditions, the number of latent variables that the method identified (i.e., number of latents returned by TETRAD and number of factors in the CFA model with the lowest AIC value) was obtained. The patterns of indicators loading together on each factor were summarized in order to count the number of errors of omission (a latent variable was not found, or a latent variable had fewer than four indicators, or it did not have the correct set of four indicators) and commission (an extra latent variable was found or a latent variable had more than four indicators, or one indicator was added to a latent variable where it did not belong according to the population model).

3.5 Summary of Methods

To populate Table 2, this research proceeded in three steps:

Step1: A program in R simulated 50 replicates at each of the three sample sizes (N=100, 300, 500) for each combination of number of factors, strengths of associations, saturations, presence/absence of weak common factors and structure. The simulation generation for these 11,400 samples is outlined in Appendix Table 1.

Step 2: The Java source code for TETRAD IV (Scheines, et al. 1998; Scheines, et al. 2005; code provided by Joseph Ramsey at Carnegie Mellon University, February 2009) was used to run BuildPureClusters on the 11,400 samples. A SAS 9.2 (SAS Institute, 2008) macro was used to run CEFA 3.02 (Browne, et al. 2008) for CFA by OLS with geomin rotation, extracting 1-6 (four factor condition) or 1-8 (six factor condition) factors; AIC was computed from the $\chi 2$ and its degrees of freedom using Excel for each of those CFA output files and the solution with the smallest AIC (from the set of 1-6 or 1-8 factors) was retained. Every loading in the "winning" model was evaluated relative to its standard error to determine what indicators were assigned to what factors (ratios >1.96 were recorded as loading and ratios <1.96 were recorded as not loading). Finally, Mplus 5.2 (Muthén & Muthén, 2008) was used to run 600 confirmatory factor analyses, extracting one factor, on the zero and one-factor samples (600 trials). The one-factor CFA produces the fit statistics for the null (zero factor) as well as the specified (one-factor) model. Failures for these CFA models to converge was interpreted as "insufficient evidence of a factor" (i.e., correct detection of zero factors).

Step 3. Performance was obtained as the average value (over the 50 trials) for each condition (the number of factors identified, indicators correctly associated with their respective factors). Estimates of consistency were obtained as the percent of the 50 trials where the same number of factors was obtained.

Once the results were collected from the 50 trials per condition, nonparametric ANOVAs were carried out using SPSS (16.0, SPSS Inc. Chicago, Ill).

Chapter 4: Results

Sections 4.1 and 4.3 outline challenges in the computation of the error rates from the CCA and CFA results, respectively, that were unforeseen until the results emerged. Results of the analyses are presented separately for each method in sections 4.2 (CCA) and 4.4 (CFA), and their comparisons in section 4.5. Descriptions of CCA main effects for the conditions (Figures 4-7) are followed by the main effects analyses of variance (Table 3); condition-specific results are presented in Figures 8-13. CFA results are similarly presented as main effects for the conditions (Figures 14-17), main effects analyses of variance (Table 4) and condition-specific results (Figures 18-25). Tables 5A and 5B and Figure 26 present the results by method.

4.1 Challenges in Computing and Interpreting CCA Results

4.1.1 Defining omission errors for indicators

Estimating the consistency of solutions proved to be a great challenge in this study. The solution was to compute omission and commission errors for indicators for each factor in the CCA results separately, because the algorithm could have identified the factor comprised of items 5-8 ("factor 2") first, and that comprised of items 1-4 ("factor 1") third, fourth, or not at all. It was an insurmountable challenge to automate a complete-factor-examination for the CCA results, whereby all factors could be considered at once. Factors were considered one at a time and the CCA results were extracted manually. The variety of solutions obtained from TETRAD necessitated specification, usually based on an arbitrary-but consistently applied- decision. The complete set of error counting rules is included in Appendix 2.

4.1.2 Counting omission errors for indicators from “extra” factors

The CCA results rarely produced extra factors. However, when they did it posed a challenge for computing omission errors. As noted in introductory chapters (Chapters 1 and 2), BPC can only produce solutions consistent with simple structure: BPC can only assign an indicator to a single latent variable. This means that, whenever an extra factor was identified by TETRAD, the indicators that were assigned in that factor did not appear on any other (earlier-extracted) factor. However, it was difficult to reconcile a supernumary factor (commission error for factor) and the possibility of a not-commission error for indicator. That is, treating the extra factors as if they were *not* extraneous/supernumary for the sake of computing omission and/or commission errors for the indicators was difficult to incorporate into the scoring scheme, particularly given that the extra factors in the CFA results could be more clearly counted (as commission errors), if they duplicated already-assigned indicators. Further, it was not possible to accurately label indicators omitted from extra factors as “omissions”, since the extra factor should not have been found. It was arbitrarily decided that omission error totals for the solutions would include only omissions on the correct (or fewer) number of factors; the first 4 or 6 factors were used and the omission errors occurring in extra factors were ignored in the TETRAD results.

This was not the case for commission errors for indicators, because this type of error implies that it was extra. As outlined in the scoring rules (Appendix 2), commission errors for indicators were computed as the number of commission errors divided by the number of times that factor was found- whether the factor was less than, equal to, or greater than the correct number of factors in the population giving rise to that sample.

4.1.3 Success of the sentinels for identifying omission errors for indicators

As outlined in Chapter 3, a method for determining which factor a set of indicators was representing had to be included so that determinations of whether an indicator assignment to a factor was correct, or constituted a commission error, could be identified. Similarly, identifying omissions of indicators from factors could only be done when the factor itself was known, so that its indicators could be evaluated (for whether they had been included or omitted). The method used in this study was to treat one (the first) indicator on each factor as a sentinel, so that when that indicator was observed, the other indicators that should have loaded with it would be known. Then determination of what was (correct) or was not (commission) supposed to load with the sentinel could be made, as well as whether what should have been there was missing (omission). Sentinels were programmed in the simulation as having pathweights 0.10 stronger than the non-sentinel indicators of a factor in the population; so the strong loading condition featured sentinels loading at 0.85 and non-sentinels loading at .75 on their factor in the population. To be consistent over all error identification, characterization and counting, the sentinels were used to determine what other indicators should/should not have been assigned to the same factor. Appendix 2 outlines the rules for assigning errors when more than one (or two or three) sentinels were assigned to the same factor. All scoring decisions were made for consistency and interpretability of the error rates –within the CCA results and also for use with the CFA results.

4.1.4 Observing one-indicator factors identified by vanishing tetrads

The results for some conditions were unexpected in the sense that vanishing tetrads, and not-vanishing partial correlations, between some subset of four indicators resulted in sufficient evidence for a common, latent cause to be identified: that is, the result returned a factor. However, although the evidence was sufficient across the set of four indicators to support a factor, only one indicator was assigned to the factor. This apparent contradiction was rarely observed in some conditions but in others, nearly all of the 50 trials reflected this outcome. The developer of the particular algorithm used (BuildPureClusters) specified that such outcomes are possible, and while they provide little or no information about the structural model, the measurement model-specific information they provide can be used and interpreted (R. Silva, personal communication 29 April 2009). When this was observed, omitted indicators were scored as omission errors, and the conditions with these outcomes simply tended to have higher omission error rates.

4.1.5 Error rates for indicators increased as factors were ‘found’

The CCA results were analyzed by factor (i.e., what was observed in the first factor found, what was observed in the second factor found, etc.). As might be expected given the nature of the algorithm, and the fact that it proceeds sequentially, not simultaneously like the CFA estimation, the error rates were not constant across factors but increased. Total error rates were computed as the sums (omission, commission) over the entire solution, rather than a per-factor average or other type of summary.

4.2 CCA Results

4.2.1 Discovering dimensions, missing or mis-assigning indicators: CCA

Given the foregoing, and together with the rules articulated in Appendix 2, the results for CCA at uncovering the dimensionality of the data can be discussed. Descriptions of results (Figures 4-7) are followed by the main effects ANOVA results (Table 3) and condition-specific results are shown in Figures 8-13.

CCA returned zero factors for every sample generated from uncorrelated indicators (100% accurate and 100% consistent). Collapsing across sample sizes and loadings, CCA found an average of 0.93 (SD 0.05) factors in samples generated from one-factor populations was. Any result that identified a factor was 100% correct at assigning all four indicators to that factor (commission errors were not possible); the average omission rate for one-factor samples was 0.28 (SD 0.21). That is, on average, 0.28 indicators were left out of these four-indicator solutions; all results in one-factor data either assigned all four indicators or zero.

Four factor samples yielded an average of 2.86 (SD 1.16) factors found, and collapsing over all conditions (apart from number of latents in the population) an average of 1.14 (SD 1.16) latent variables were omitted from each solution. The average omission error rate for indicators (including four indicator omissions for every missed latent), was 1.70 (SD 0.79). In other words, for every result based on samples from a four-factor population, an average of 1.1 latent variables, and 1.7

indicators were omitted. An average of 0.03 (SD 0.01) extra factors were found and 0.48 (SD 0.64) indicators were mis-assigned in four factor data.

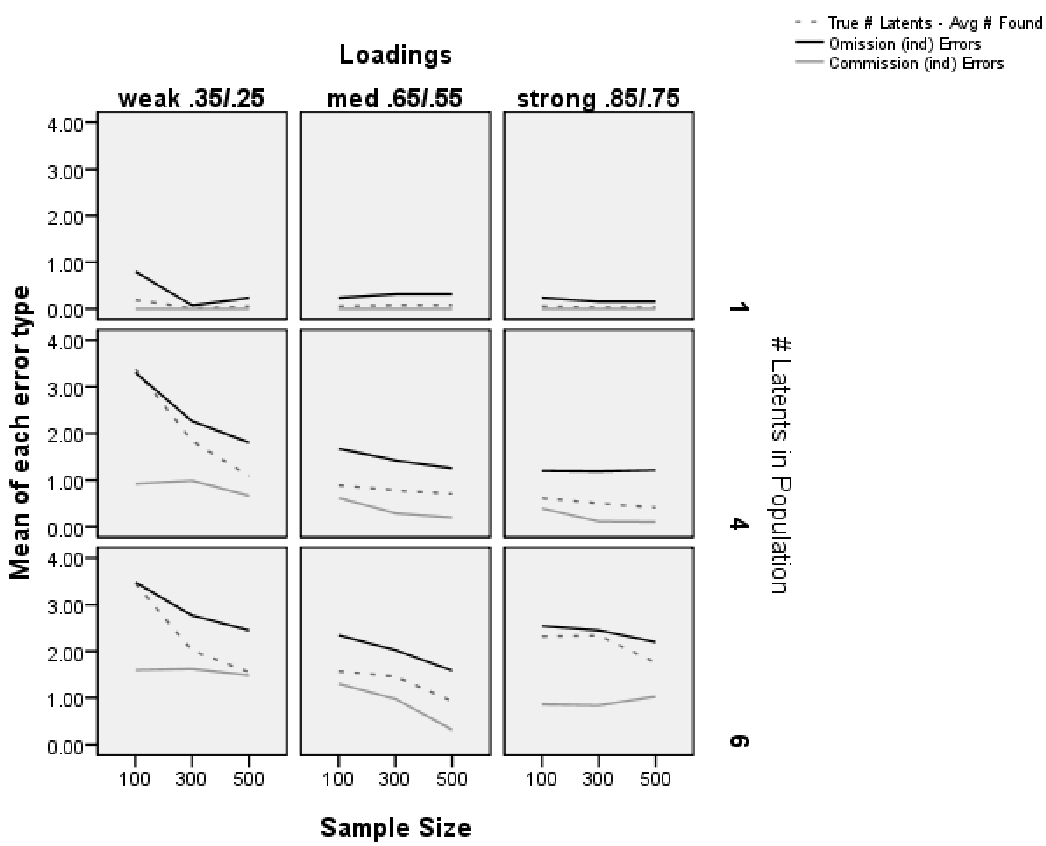
Six factor samples yielded an average of 4.07 (SD 1.72) factors found, with an average of 1.93 (SD 1.84) latents omitted and 0.06 (SD 0.06) extra factors found. An average of 2.4 (SD 1.34) indicators were omitted from six factor solutions and 1.11 (SD 1.69) indicators were mis-assigned. In the plots that follow, commission of latents (finding extra factors) errors are not included since they were so low when averaged over conditions; they are included in the one-way (nonparametric) ANOVAs shown in Table 3. The plots that follow focus on omitted latents, and indicator omissions and commissions.

The loadings condition represented an important test for TETRAD since performance is based on proportional, and not absolute values of the correlations. Figure 4 shows the error rates (average number of omitted factors, indicators mis-assigned (commission errors) and indicators un-assigned (omission errors)) by sample size for the three loadings conditions, collapsing over purity, presence of a weak factor, and whether factors were correlated or orthogonal.

The one-factor results appear in the top row of plots. For these samples, omission and commission error rates are overlaid around zero, as no commission errors were made and the omission errors were low due to the accuracy (average # latents omitted, grey dashed line). Unlike the one factor results, the omitted factor fate (grey dashed line) for four and six factor tended to be flat or decrease as sample size increased, although for six factors with strong loadings, sample size of 500 (largest) seemed to decrease performance/increase factor omission rate slightly. All three types of errors seemed to vary with the loadings (columns).

In general, results for the medium and strong loadings are similar within four factor results; omission errors were higher when loadings were strong in six factor data as opposed to when loadings were medium sized; commission errors were higher and flat when loadings were weak in six factor data, tended to decrease with sample size when loadings were medium, and were flat or tending upwards as sample size increased when loadings were strong.

Figure 4. CCA Error rates by sample size and loadings for one, four and six factors.



Comparing the omission and commission error rates for four and six factors across each loading level, the patterns seem similar. As expected given the lower bound for detectible correlations in TETRAD (given by Spirtes, et al. 2000) of 0.30, higher errors were observed for the weak loadings condition; within four and six factor data, the effects of loadings are less pronounced. It is possible that CCA performance was more susceptible to experimental conditions when there were more factors and stronger loadings; that is, weak factor loadings were only above .3 when loadings were strong. Some errors made in the strong-loading condition might be most easily attributable to the conditions, whereas at least some of the errors in lower-loadings conditions must be attributable in part to the threshold of TETRAD's sensitivity.

Purity of factors (whether two latents shared indicators in the population), presence of a weak factor, and independence of factors were only features in the four- and six-factor samples. Figure 5 represents the error rates (omitted factors, indicators mis-assigned (commission errors) and indicators un-assigned (omission errors)) by sample size for the three purity conditions, collapsing over loadings, presence of a weak factor, and whether factors were correlated or orthogonal. In Figure 5 it can be seen that errors tended to decrease as sample size increased, and this was true for all purity conditions; the accuracy (omitted factors, grey dashed line) tended to increase with sample size but decrease with decreasing purity. This sensitivity to purity was more pronounced in the six-factor results, with the four factor share 0 and share 1 plots more similar and the share 2 plot reflecting mostly an increase in factor omission errors. The effect of purity on indicator omission rates (black solid line) did not appear to be marked for four factor data, but for six factor data, the indicator omission error rates increase as purity decreased (went from pure (share 0) to two latents sharing 1 to sharing 2 indicators).

Errors of commission for indicators (solid line) were generally low in the four factor condition and were higher in the six factor condition.

Figure 5. CCA Error rates by sample size and purity for four and six factors.

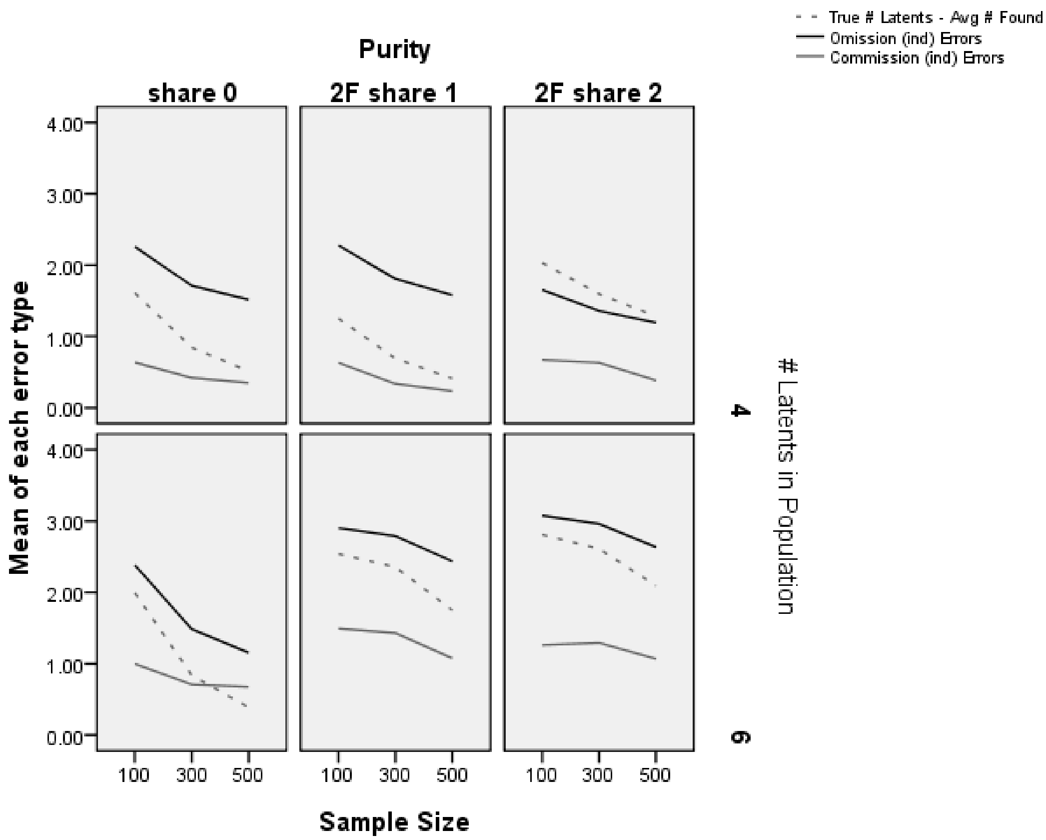
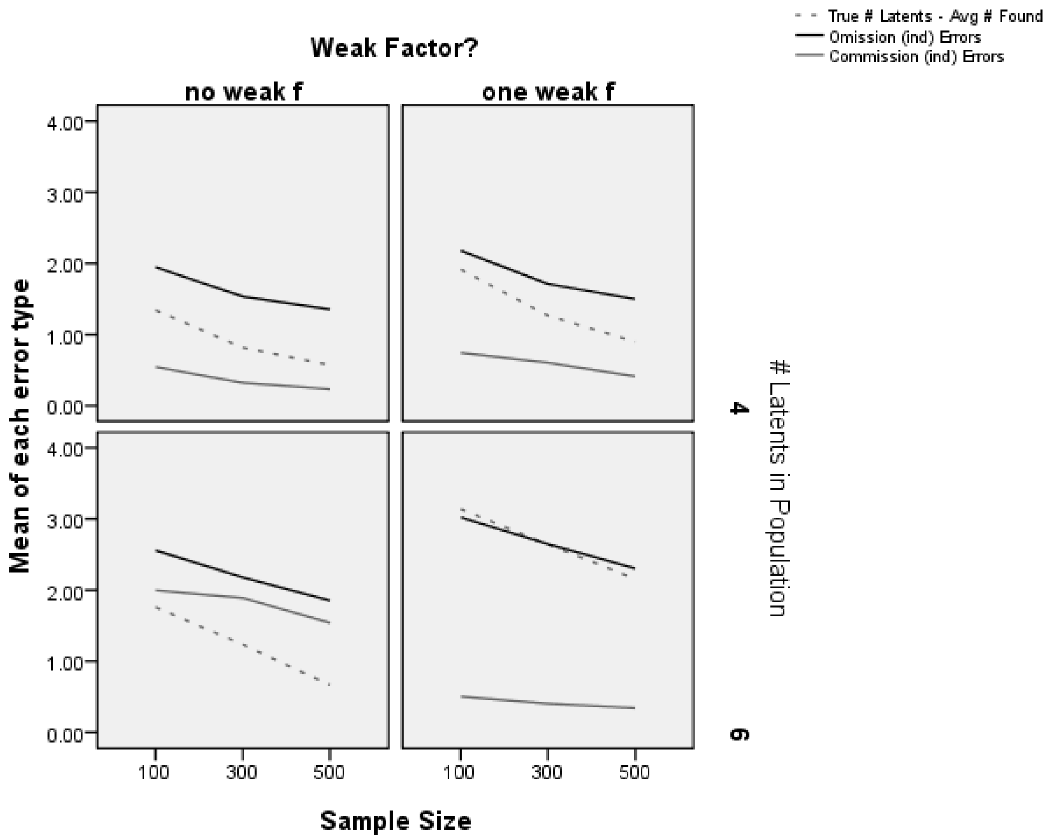


Figure 6 represents the error rates by sample size for the conditions with and without a single weak factor, collapsing over purity, loadings and whether factors were correlated or orthogonal.

Figure 6 suggests that the presence of a weak factor decreased dimensional discovery rates (factor omission, grey dashed line, increases), although factor discovery improved (omission errors decreased) as sample size increases. The omission and commission error rates are closest (in four plots) for six factor data with no weak factor, while with one weak factor the commission errors for indicators is very low and omission errors for indicators are fairly high (the highest of all four plots).

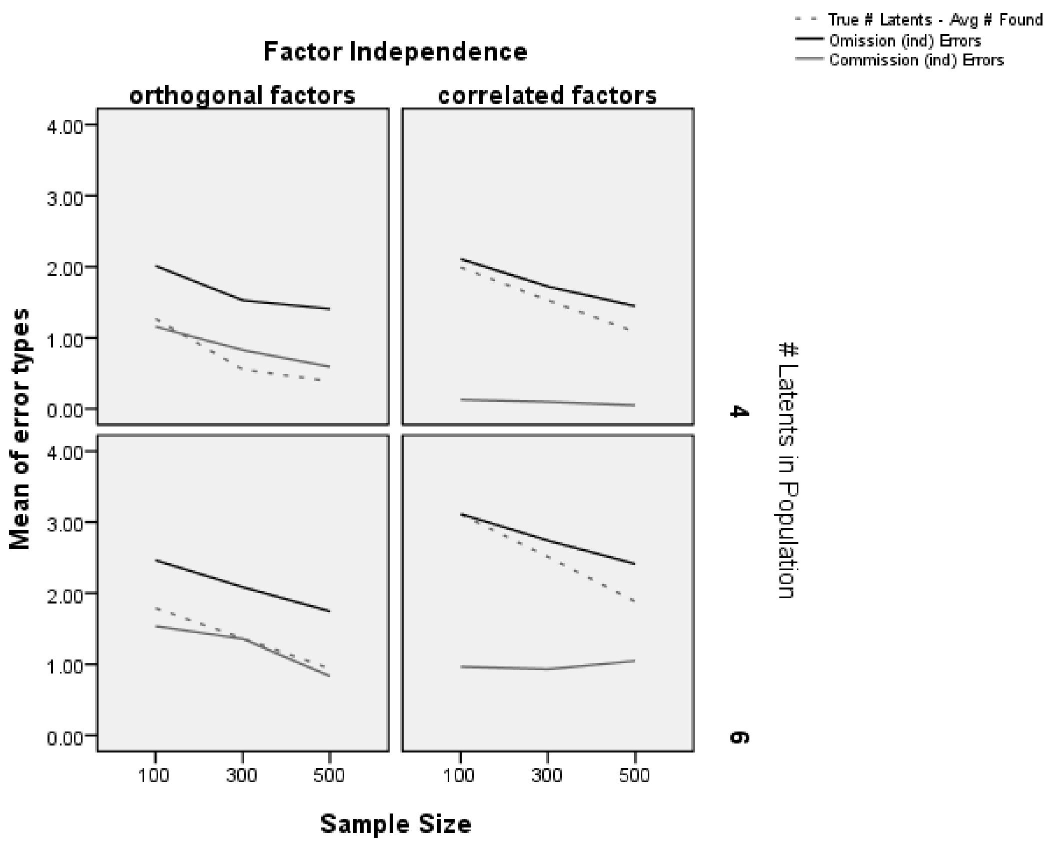
Figure 6. CCA Error rates by sample size and presence of weak factor for four and six factors.



The effect of a weak factor on indicator omission rates (black line) is similar for four and six factor data, namely, this error rate was slightly higher if a weak factor was present. Errors of commission for indicators (grey line) increased slightly for four factors, while clearly decreasing for six factors, when a weak factor was present. The effect of a weak factor might be greater for indicators and their correct assignment than for the discovery of latent variables, although in the six factor condition, the factor omission error is increased in the presence of a weak factor.

Figure 7 represents the error rates by sample size for the conditions when factors were orthogonal or correlated at ρ=.5, collapsing over purity, loadings and whether there was a weak factor.

Figure 7. CCA Error rates by sample size and independence for four and six factors.



As was observed with the addition of a weak factor (Figure 6), when factors were correlated, the omission of factors (dashed line) and omission of indicators (black line) rates were worse than when factors were orthogonal. This effect appears most striking for the six factor data (lower right plot). Since recovery of factors was worse when factors were correlated, the lower commission error rates (which could only be committed if a factor was found) are also lower when factors are correlated.

As will be seen in Figures 8-13, the variances of values for the main effects were not necessarily equal across groupings, so one-way nonparametric ANOVAs

(Kruskal-Wallis) were carried out comparing the three main error rates (factor and indicator omission, indicator commission) across the levels of the condition variables. These results are given in Table 3.

Table 3. Main effects for conditions on errors from CCA

	Omission errors- Factors	Commission errors- factors	Omission errors- indicators	Commission errors - indicators
Sample size (100, 300, 500)	χ^2_2= 9.31 **	χ^2_2=3.53 p=0.17	χ^2_2=8.08 *	χ^2_2= 6.71 *
Number of factors (0, 1, 4, 6)	χ^2_2= 28.20 ***	χ^2_2= 0.10 p=.75	χ^2_2= 45.07 ***	χ^2_2= 25.79 ***
Strength of relations (.8, .6, .4)	χ^2_2=30.33 ***	χ^2_2=0.67 p=.41	χ^2_2= 37.60 ***	χ^2_2= 8.70 *
Saturation (pure, share 1, share 2)	χ^2_2= 30.70 ***	χ^2_2= 1.26 p=.53	χ^2_2= 12.72 *	χ^2_2= 10.31 *
Weak factor (yes, no)	χ^2_1= 20.64 ***	χ^2_1= 3.33 p=0.068	χ^2_1= 5.62 *	χ^2_1= 1.73 p=0.19
Structural model (independent, correlated factors)	χ^2_1= 22.76 ***	t(7)=2.96 *†	χ^2_1= 2.37 p=0.12	χ^2_1= 39.68 ***
* p<0.05; ** p<0.01; *** p<0.0001 † one-sample t-test comparing rate in indedpendent factors condition to zero.				

The results in Table 3 suggest that all conditions affected CCA's performance at dimensional discovery in terms of omitting factors from the solution when collapsing across the other conditions –including number of factors in the population. None of the conditions affected the tendency to commit factor errors (find more factors than there were in the population) except for the presence of correlated factors: none of the commission of factors errors occurred when factors were orthogonal in the population. The 0.03 extra factors found, on average from four factor and 0.06 extra factors found, on average from six factor samples were all errors made when population factors were correlated at 0.5. These could not be compared across the conditions so the values were instead compared by one-sample t-test against the reference value zero; significantly more commission errors were committed when factors were orthogonal than when factors were correlated, where zero commission errors occurred.

4.2.2 CCA performance by conditions

Figure 8 presents CCA performance at dimension discovery (factor omission errors) for the four and six factor data across loadings and whether factors were orthogonal. Within each panel, the performance under the three purity conditions (share 0, open diamond; 2 latents share 1, star; 2 latents share 2, filled circle). As noted, these conditions were not part of the zero and one factor results. There are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and presence of a weak factor.

From the slopes of the lines in each plot shown in Figure 8, it can be seen that, when factors are pure, factor omission rates decrease (accuracy increases) as loadings increase (dashed line); this is true whether factors are orthogonal or correlated. That is, the grey solid line slopes downward in each plot. When factors are orthogonal (top left plot), accuracy increases with loadings, and performance with pure factors in the four factor data is between performances where two latents share one (better) and two (worse) indicators. For four factor data when factors are correlated at .5, sharing 0 and sharing 1 indicator conditions did not dramatically affect performance. However, in the six factor data, sharing had a dramatic effect on accuracy and this did not seem to matter whether the factors were correlated or orthogonal. In the six factor data, it seems that effects of loadings to decrease factor omission errors are undone if impure factors are present. This is more pronounced than in the four factor results.

Figure 8. CCA dimension discovery by loadings/independence conditions, four and six factors

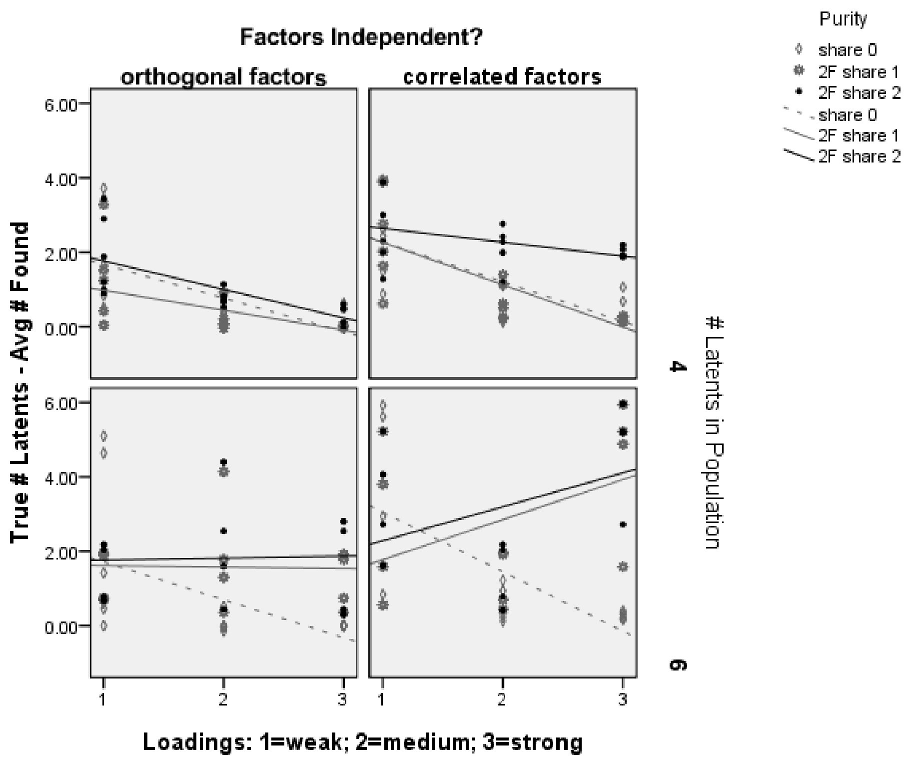
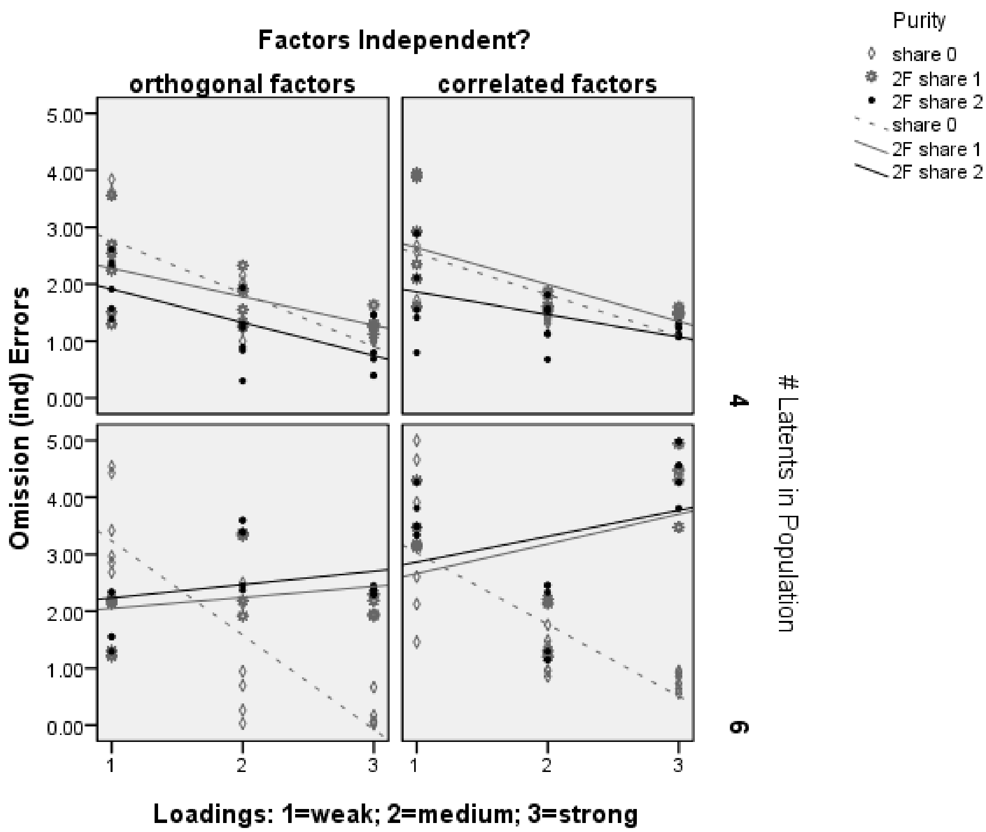


Figure 9 shows the omission error rate (for indicators) across conditions for the four and six factor data. The indicator omission rates decrease as loading strength increases for both four and six factor data with pure factors, but the relationship to loadings tends to be stronger in the six, relative to the four, factor data. When any sharing is present (solid lines), omission errors tend to increase as loadings increase for six factor data, and not for four factor data.

Figure 9. CCA Omitted indicators per solution across loadings/independence conditions for four and six factors.

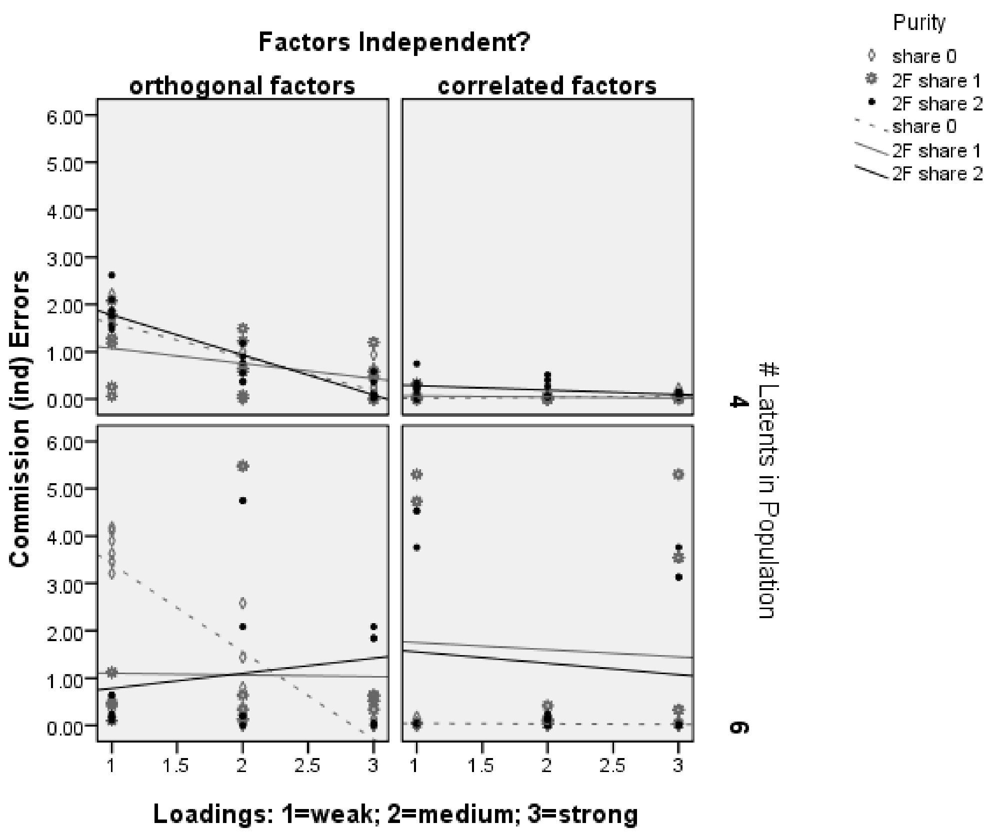


As with Figure 8, there are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and presence of a weak factor. The variability in these six points shrinks as loadings increase, and this trend is observed for all purity condition levels.

Similar to the factor omission error results (Figure 8), omission errors for indicators tended to decrease (i.e., CCA performance improved) with strength of loading for both four and six factor solutions. Only the six factor results reflect a sensitivity to shared indicators in terms of indicator omission error rates.

The errors of commission for indicators, where the CCA solution placed an indicator away from its sentinel or either with or without its sentinel but in the minority on a factor, are shown in Figure 10. The commission error rates were fairly small, reflecting the conservative nature of TETRAD (i.e., the algorithms will omit a path when there is evidence of no path as well as when there is insufficient evidence *for* the path; an extra path is less likely by design).

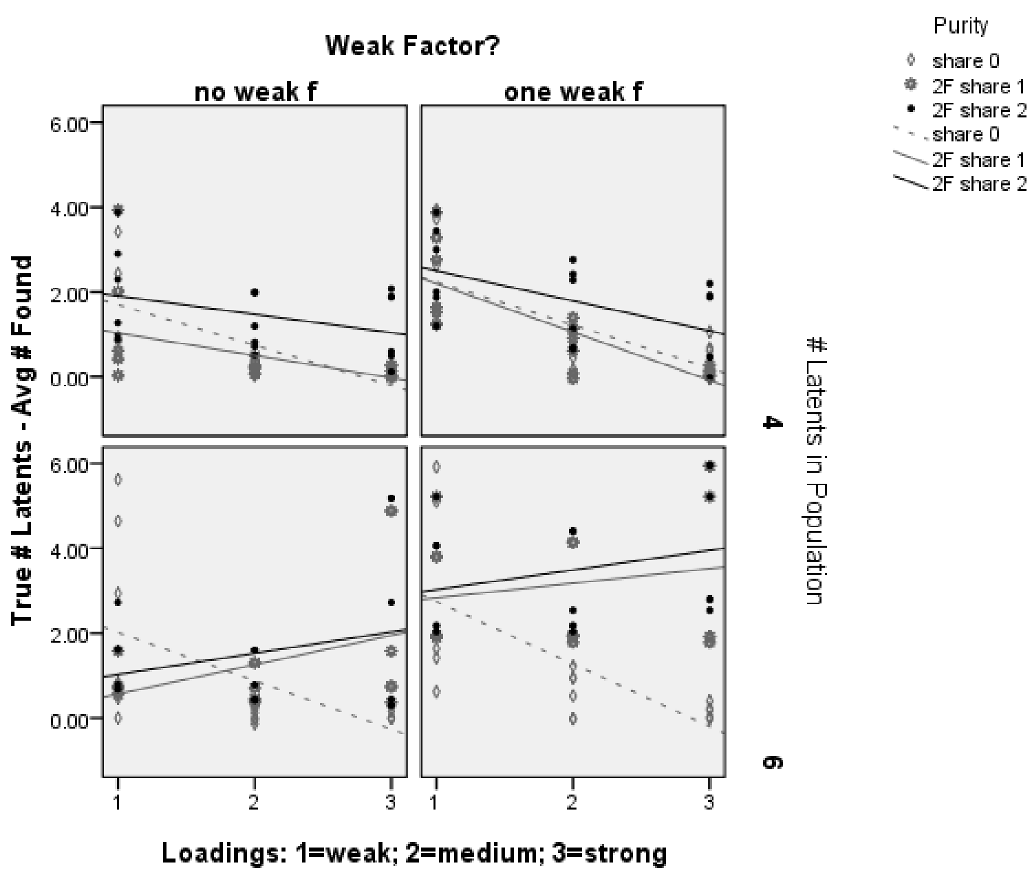
Figure 10. CCA committed indicators per solution across loadings/independence conditions for four and six factors.



When factors were correlated, as reflected in the accuracy plots shown in Figure 8, fewer factors were found; thus lower commission error rates for the correlated factors conditions (right plots) is expected. Like omission errors, sharing one or two indicators did not alter error rates given that the factors were impure (solid lines are very close in all conditions). Unlike the omission errors, however, Figure 10 shows that only commission errors for six orthogonal factors (lower left plot) decreased with loading strength; commission errors were essentially flat in all other contexts.

Figure 11 presents CCA performance at dimension discovery (factor omission errors) for the four and six factor data across loadings and the presence of a weak factor.

Figure 11. CCA dimension discovery by loadings/weak factor conditions, four and six factors



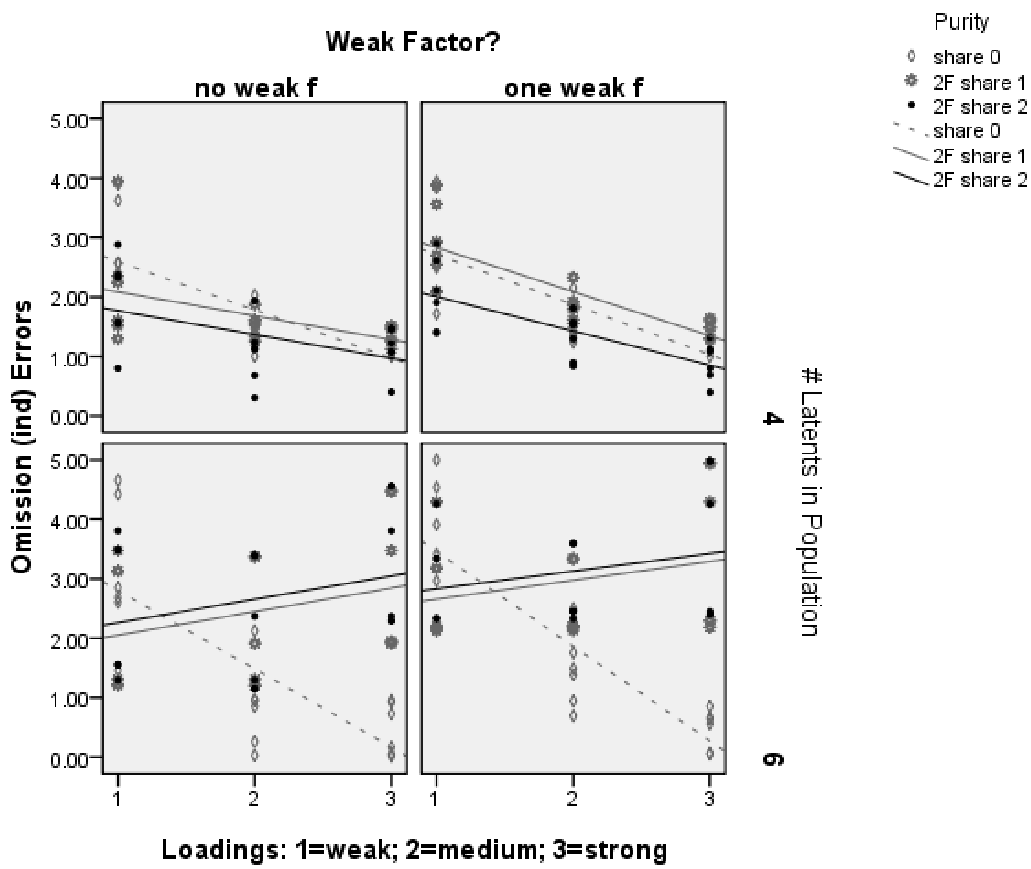
Within each panel, the performance under the three purity conditions. As noted, these conditions were not part of the zero and one factor results. There are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and whether factors were orthogonal.

From the slopes of the lines in each plot shown in Figure 11, it can be seen that, when factors are pure, factor omission rates decrease (accuracy increases) as loadings increase (dashed line); this is true whether there is or is no weak factor. That is, the grey solid line slopes downward in each plot.

When all factors have the same loading strength (top left plot), accuracy increases with loadings, and performance with pure factors in the four factor data is between performances where two latents share one (better) and two (worse) indicators. For four factor data in the presence of a weak factor, sharing 0 and sharing 1 indicator conditions did not dramatically affect performance. However, in the six factor data, sharing had a dramatic effect on accuracy and this did not seem to matter whether there was a weak factor. In the six factor data, it seems that effects of loadings to decrease factor omission errors are undone if impure factors are present. This is more pronounced than in the four factor results. The effects of the weak factor are similar to those of the independence of factors (Figure 8).

Figure 12 shows the omission error rate (for indicators) across conditions for the four and six factor data. The indicator omission rates decrease as loading strength increases for both four and six factor data with pure factors, but the relationship to loadings tends to be stronger in the six, relative to the four, factor data. When any sharing is present (solid lines), omission errors tend to increase as loadings increase for six factor data, and not for four factor data.

Figure 12. CCA Omitted indicators per solution across loadings/weak factor conditions for four and six factors.



As with Figures 8-11, there are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and whether factors were orthogonal.

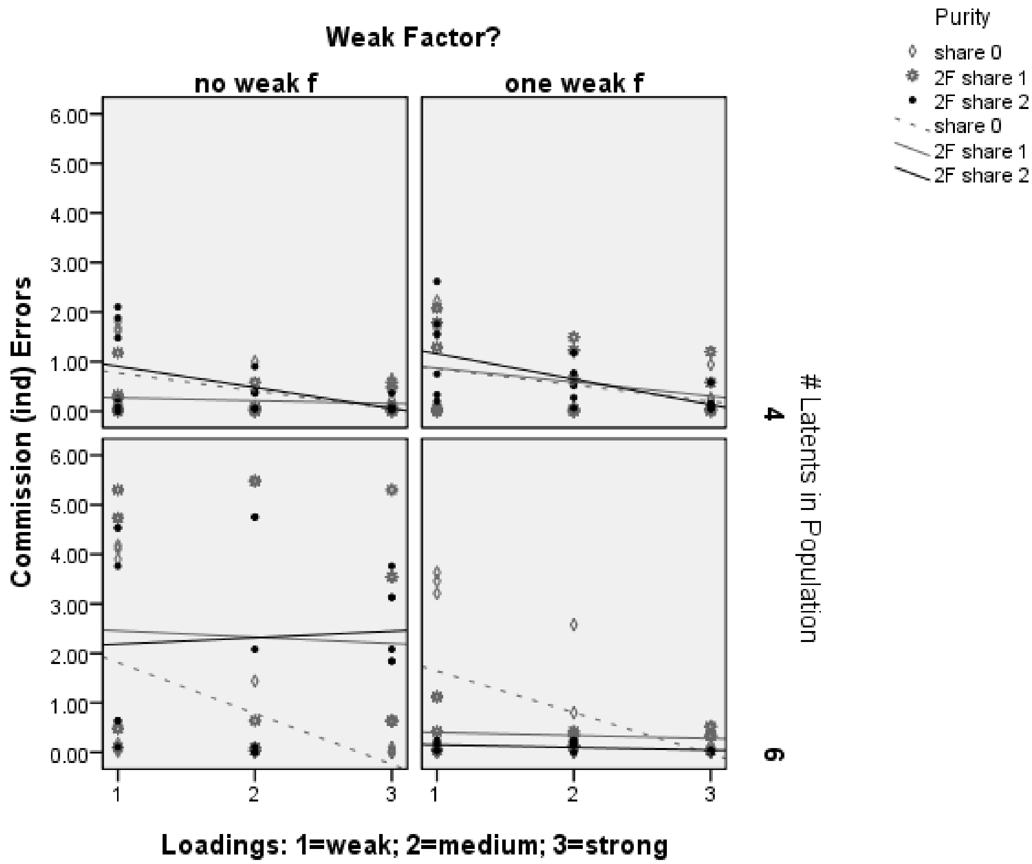
Similar to the factor omission error results (Figure 11), omission errors for indicators tended to decrease (i.e., CCA performance improved) with strength of loading for both four and six factor solutions. Only the six factor results reflect a sensitivity to shared indicators in terms of indicator omission error rates. These results stratified by presence of a weak factor are similar to those shown in Figure 9 stratified by independence of factors.

The errors of commission for indicators, where the CCA solution placed an indicator away from its sentinel or either with or without its sentinel but in the minority on a factor, are shown in Figure 13. The commission error rates were fairly small in CCA solutions, reflecting the conservative nature of TETRAD (i.e., the algorithms will omit a path when there is evidence of no path as well as when there is insufficient evidence *for* the path; an extra path is less likely by design). Since CCA performance at finding factors was lower when weak factors were present, commission of indicator errors were also lower (right plots, Figure 13) for four and six factor data.

Like indicator omission errors, sharing one or two indicators did not alter commission of indicator error rates in four factor data given that the factors were impure (solid lines are very close in all conditions). Unlike the omission errors shown in Figure 12, only commission errors for six orthogonal factors (lower left plot) decreased with loading strength; commission errors were essentially flat in all other contexts.

The presence of a weak factor has no apparent impact on the commission of indicator error rates in the four factor data (top plots) but the overall accuracy being lower in six factor data when a weak factor is present, relative to when there is no weak factor, results in a drop in the commission error rates when factors are impure (solid lines, lower two plots).

Figure 13. CCA committed indicators per solution across loading/weak factor conditions for four and six factors.



Overall, these results for the CCA analyses have emphasized accuracy, since invariance or consistency, defined as (yes/no) observing 95% of the same solution (number of latent variables, whether accurate or not) within each set of 50 analyses, was only observed for the zero and one factor results. The commission error rates were fairly small, reflecting the conservative nature of TETRAD (i.e., making extra factors and paths less likely in CCA solutions), however the effects of the conditions was still significant for commission of indicator errors. The omission of factor and omission and commission of indicator errors were all significantly impacted by these conditions in CCA results. The commission of factor errors was only significantly decreased by correlation among the factors, which might be due to lower levels of factors found in that condition (so commission of factors was very unlikely).

4.3 Challenges in Computing and Interpreting CFA Results

4.3.1 Identifying loaders when pathweights were not significant

As described in Chapter 3, the AIC values for models were used to select the CEFA solution with the best fitting number of factors, i.e., to determine the number of found factors in the factor analytic solutions. Confirmatory factor analyses, fitting a one-factor model, were used to determine whether a zero factor or a one factor solution fit the data best. The confirmatory analysis output from Mplus includes a test of the null (zero factor) model, streamlining these 600 analyses. Failures of these models to converge (after 2000 iterations) were observed 10%-40% of the time in the zero factor samples. When the models did converge, in every case AIC supported a one factor model over the zero factor (null) model, but the estimated loadings of the indicators were not significantly different from zero – the individual item loadings were not significant, but their magnitudes and pattern suggested evidence for a common factor. In all but two cases where AIC supported a one-factor model from a zero-factor sample, no factor loadings were estimated at all.

Thus, the common factor analysis results included inferred zero factor solutions when the confirmatory factor analysis with all 4 indicators loading on one latent variable failed to converge (this only occurred for true zero factor samples); supported one-factor solutions when AIC supported a one factor model over the null (zero factor) model, but with no factor loadings estimated; and supported one-factor solutions when AIC supported a one factor model over the null model and with factor loadings estimated and, relative to the loading standard errors, loadings were either significant (loading/SE≥1.96) – resulting in the indicator being "assigned" to that factor, or not significant (loading/SE<1.96) –resulting in the indicator not being assigned to that factor.

4.3.2 Defining and counting omission errors for indicators from "extra" factors

Estimating the consistency of solutions in the CFA programs created challenges for the estimation of consistency of solutions from the factor analyses as well, because it was important that the results from the two methods were comparable in some senses. The error counting rule set provided in Appendix 2 was applied to the CFA results. The scoring rules can be summarized as comprising two stages and two defaults. Stage 1: look for a sentinel, the variable of the factor in the population that signals which population factor has been found. Stage 2: if no sentinel is found, the found factor is identified as the one in the population to which the majority of assigned indicators belong. Default 1: if the sentinel and majority rules do not lead to the identity of the found factor, identify the found factor as that to which the first indicator(s) belong; all other indicators (missing or present) are errors, classified as appropriate. Default 2: for shared indicators conditions only, if there is no sentinel for population factors 2 (x5) or 3 (x9) and x8 (S1) and/or x7 (S2) appear, treat them as if they represent factor 2.

The CFA results produced extra factors more often than the CCA results did, with the additional complication that CFA results could have indicators loading twice in a solution. As noted elsewhere, while the BuildPureClusters algorithm of

TETRAD's correlation constraint analysis (BPC) can only assign an indicator to a single latent variable, this is true of neither confirmatory factor analysis (for the zero and 1- factor solutions) nor common factor analysis (for the 4- and 6- factor solutions).

4.3.3 Success of the sentinels for identifying omission errors for indicators

The use of sentinels in identifying factors for the purpose of identifying, classifying and counting errors led to different difficulties in the CFA context. While sentinels were useful in decisionmaking in TETRAD solutions, since sentinels and their non-sentinel co-loaders could appear on >1 factor, this complicated the error counting. That is, it was not possible to observe one sentinel loading on two factors in CCA but it *was* possible, and was observed in the CFA results. The CCA errors were achieved factor by factor, so that commissions like one indicator loading twice within one solution were not specifically counted –they were treated as simple indicator commission errors.

4.3.2 Observing one-indicator and zero-indicator factors

Because we considered indicators to load on factors in the CFA results only if the pathweight divided by its standard error was greater than 1.96 (statistically significantly different from zero- with no consideration for adjustments for the multitudes of comparisons), it was possible for one and even zero indicators to be assigned to a factor that the solution found.

4.4 CFA Results

4.4.1 Discovering dimensions, missing or mis-assigning indicators: CFA

Given the foregoing, and together with the rules articulated in Appendix 2, the results for CFA at uncovering the dimensionality of the data can be discussed. Descriptions of results (Figures 14-17) are followed by the main effects ANOVA results (Table 4) and condition specific results (Figures 18-25).

As described earlier, confirmatory factor analysis (with maximum likelihood estimation) was used to fit a one factor model to the zero and one factor data in order to obtain the AIC values that would be used to identify the best fitting factor model in the four and six factor data. For the zero and one factor data, the confirmatory analyses proceeded in the same manner: one factor, with variance fixed at 1, was modeled as the single cause of the four indicators in the sample. The paths of each were freed for estimation.

The confirmatory analysis results found an average of 0.73 (SD 0.09) factors in the samples generated from uncorrelated indicators. This average represents the fact that 72.67% of trials with zero factor data supported (by AIC) the one-factor model over the null model; in the 27.3% of cases where the one-factor model was not found to fit better than the null, the model failed to converge, which was interpreted as a zero-factor model (i.e., no model) fitting better. Although 72% of the analyses reflected a better fit for one factor, only two of 150 trials yielded any indicators with paths that were significantly different from zero (path/SE(path) > 1.96). Thus, given that 108 of 150 trials found a factor and only a total of five indicator commission errors were made for the zero factor data, the commission error rate for indicators was 4.6%.

When compared to the null model, the one-factor confirmatory factor analysis models fit to the one-factor samples never supported the null model (100% accurate and 100% consistent). Every result in one-factor data assigned all four indicators and paths were always statistically significantly greater than zero, so neither commission nor omission errors were observed for one-factor data.

The four- and six-factor data were analyzed using common factor analysis and the number of factors in the model with the best AIC (of all models fit to each sample extracting from 1 to correct # +2 factors) was recorded as the "winning" fit.

Based on these AIC-identified solutions, four-factor samples yielded an average of 3.72 (SD .75) factors found, and collapsing over all conditions (apart from number of latents in the population), with an average of 0.28 (SD 0.75) latent variables omitted from each solution. The average omission error rate for indicators (including four indicator omissions for every missed latent), was 1.22 (SD 0.50). In other words, for every result based on samples from a four-factor population, an average of 0.28 latent variables, and 1.2 indicators were omitted. An average of 0.18 (SD 0.07) extra factors were found and 4.32 (SD 1.94) indicators were mis-assigned. Six factor samples yielded an average of 5.13 (SD 1.55) factors found, with 0.87 (SD 0.1.55) extra factors found. An average of 1.13 (SD .50) indicators were omitted from six factor solutions and 8.08 (SD 3.45) indicators mis-assigned. Unlike for the CCA results, commission of latent variable errors were anticipated to be sufficient for analysis (e.g., Silva 2005 reported them be the more frequently occurring error for factor analysis in his simulation) so unlike the CCA results, they are included in the CFA plots and analyses that follow. (In section 4.5, when the results from the two methods are compared, commission of latent variable errors will be considered from both methods).

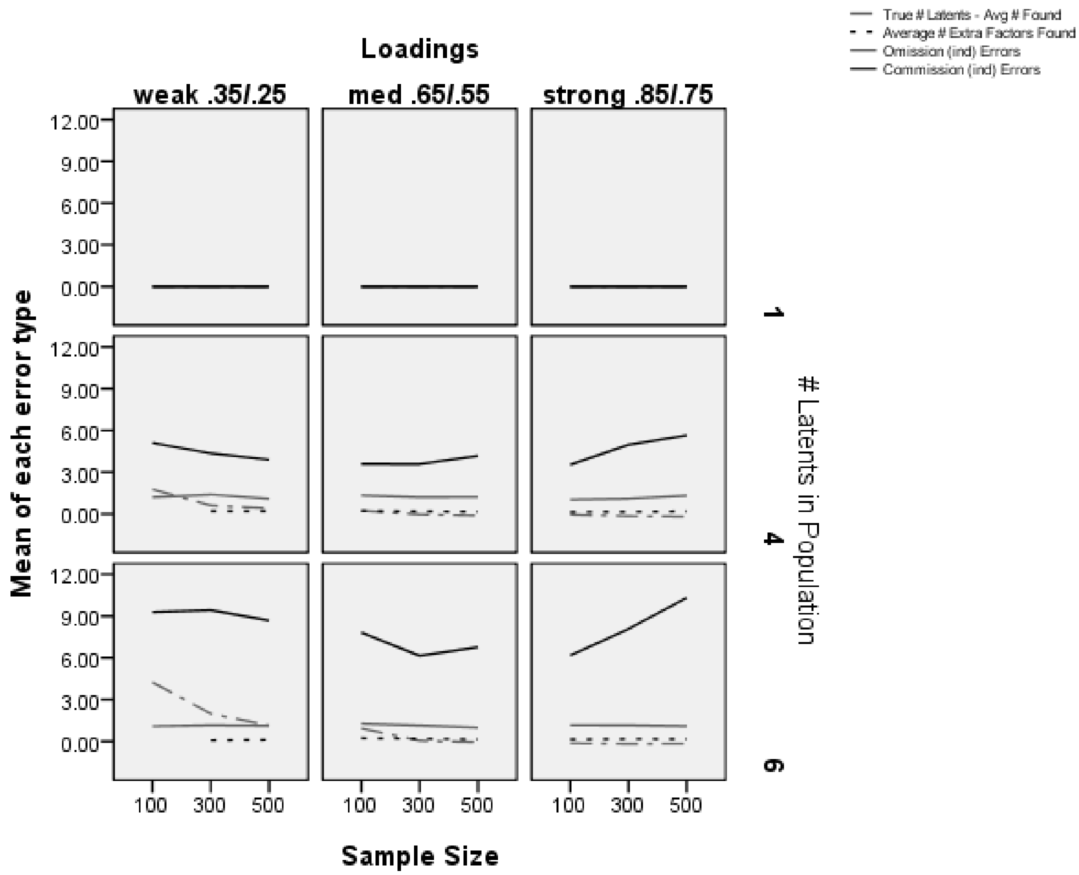
Figure 14 shows the error rates (average number of omitted factors, indicators mis-assigned (commission errors) and indicators un-assigned (omission errors)) by sample size for the three loadings conditions, collapsing over purity, presence of a weak factor, and whether factors were correlated or orthogonal.

The one-factor results appear in the top row of plots. For these samples, error rates are overlaid at zero, as no errors were made on one factor data. For four factor data, commission errors for indicators appears to increase slightly as loadings increase while the other error type tended to be flat as sample size and loadings increase.

For six factor data with strong loadings, the commission error rate increased linearly with sample size; this trend was not observed in six factor data for other loadings. All four types of errors seemed to vary with the loadings in six factor data only, with the variation coming from factor omissions (dash-dot line) and indicator commission errors.

The commission of factor error rates are flat for all plots while the omission of factor error rates improve over sample size in the weak loading condition, and slightly from N=100 to N=300 in the medium loading condition, but otherwise the rate is unaffected by loading or sample size. Commission errors for indicators increased as the number of factors increased.

Figure 14. CFA Error rates by sample size and loadings for one, four and six factors.



Purity of factors (whether two latents shared indicators in the population), presence of a weak factor, and independence of factors were only features in the four- and six-factor samples. Figure 15 represents the error rates (omitted and extra factors, indicators mis-assigned (commission errors) and indicators un-assigned (omission errors)) by sample size for the three purity conditions, collapsing over loadings, presence of a weak factor, and whether factors were correlated or orthogonal. In Figure 15 it can be seen that omission, but not commission errors for factors tended to decrease as sample size increased, and this was true for all purity conditions. Omitted factor errors did not increase for share 0 to share 1 but did increase when two factors shared two indicators and this sensitivity to purity was more pronounced in the six-factor results; commission of factor rates appear unaffected by purity. Indicator errors of omission (grey solid line) and commission (black solid line) tended to increase (slightly for the omissions) as factors became less pure. The increase in indicator commission errors was fairly linear and more pronounced for six factor data than for four factor data. The factor errors were less dramatically affected by purity in either four or six factor data, but the six factor data factor omission rates decrease as sample sizes increase, and this pattern is not apparent in the four factor data.

Figure 15. CFA Error rates by sample size and purity for four and six factors.

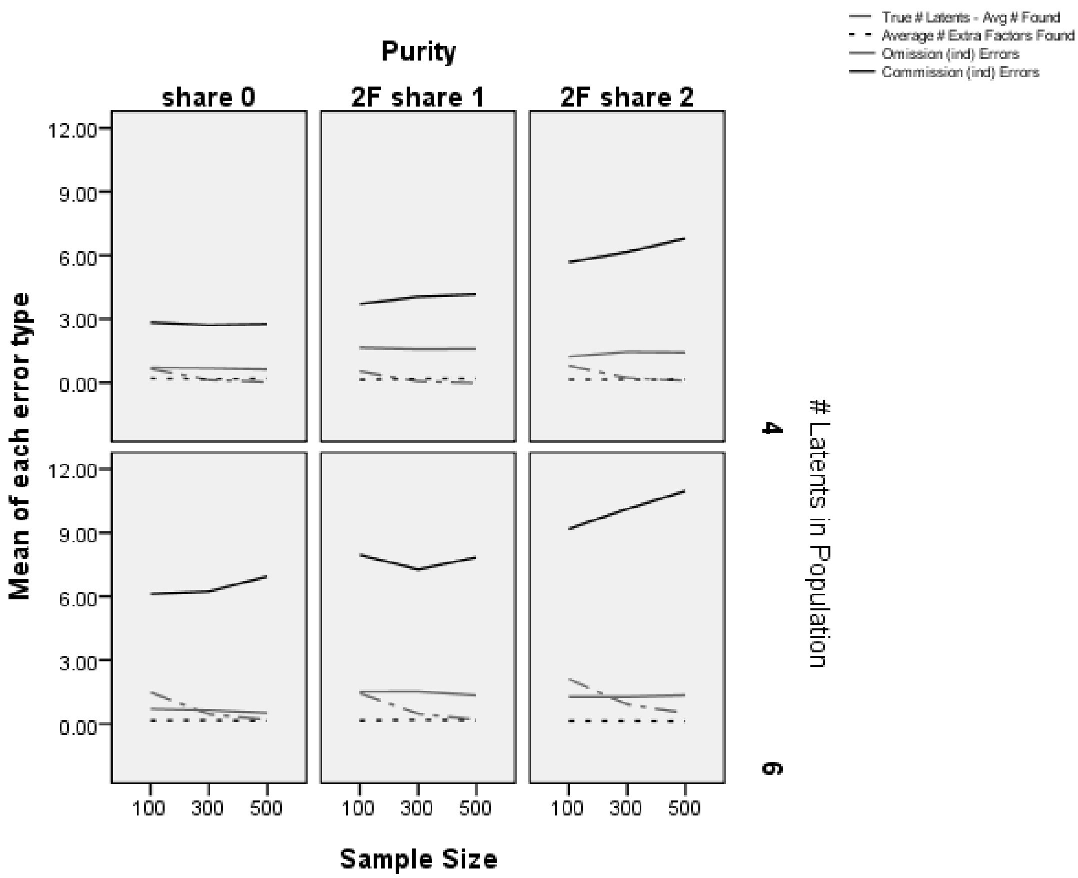
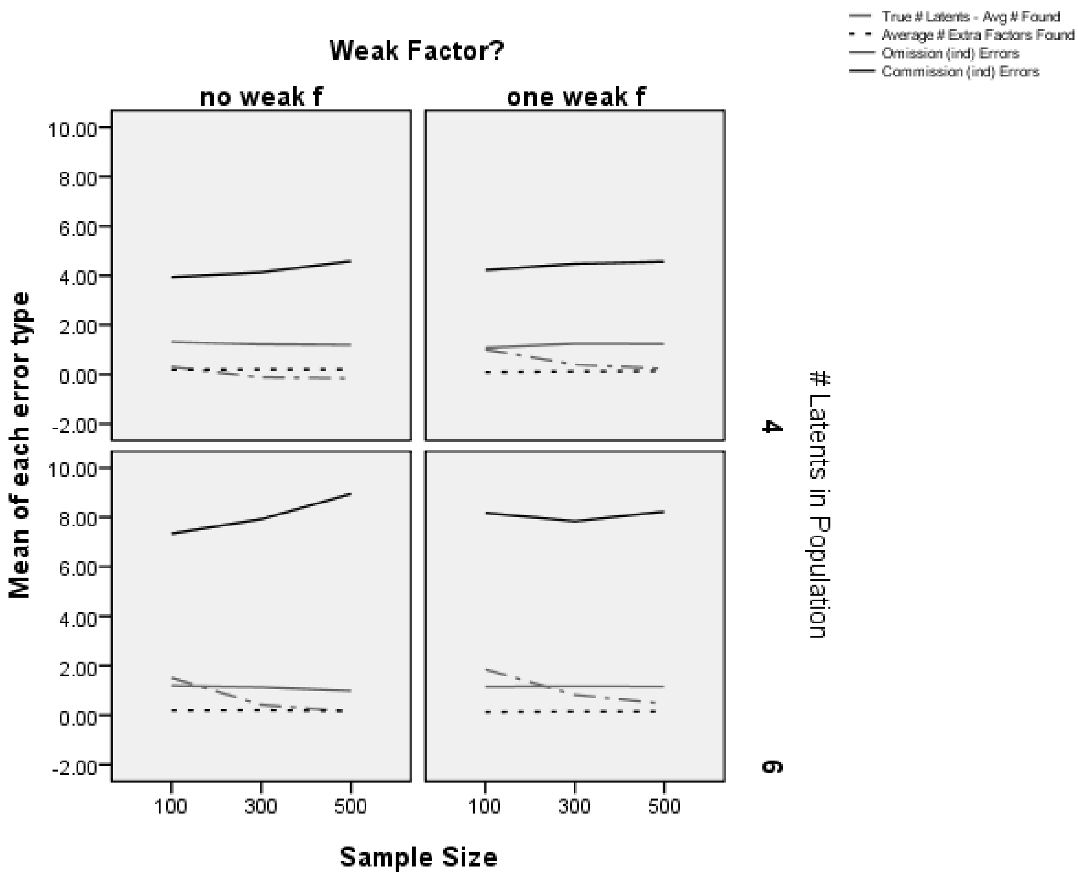


Figure 16 represents the error rates by sample size for the conditions with and without a single weak factor, collapsing over purity, loadings and whether factors were correlated or orthogonal.

Figure 16 suggests that the presence of a weak factor may slightly increase factor omission errors and slightly decrease factor commission errors, although factor discovery improved (omission errors decreased) as sample size increased. Commission error rates for indicators are highest for six factor data without a weak factor. With one weak factor the commission errors for indicators are not sensitive to sample size; increases in commission errors for indicators as sample size increases is not apparent in the presence of one weak factor for either four or six factor data..

Figure 16. CFA Error rates by sample size and presence of weak factor for four and six factors.



The effect of a weak factor is not pronounced for error rates in four factor data; the factor omission error rate (grey dash-dot line) increases, and appears to become more sensitive to sample size in the presence of a weak factor. In six factor data there is a more pronounced effect of a weak factor on the factor omission rate (grey dash-dot line) in terms of its increase in the presence of a weak factor and a greater sensitivity to sample size. The effects on indicator error rates is most pronounced for six factor data, with a sensitivity to sample size for indicator commission errors (black solid line) disappearing in the presence of a weak factor.

Figure 17 represents the error rates by sample size for the conditions when factors were orthogonal or correlated at ρ=.5, collapsing over purity, loadings and whether there was a weak factor.

Figure 17. CFA Error rates by sample size and independence for four and six factors.

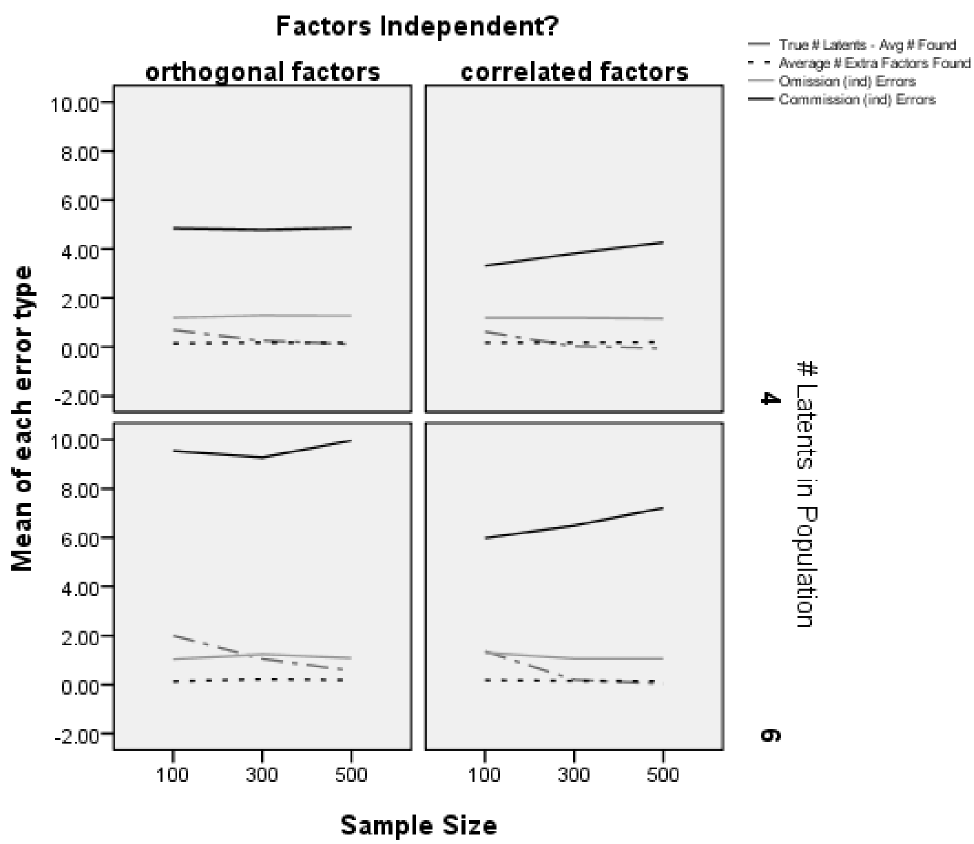


Figure 17 reflects a reduction in factor omission rates, most clear in the six factor data, as well as a reduction in the indicator omission rates for both for and six factor data. This impact of correlated factors is the opposite of this condition's effect on the CCA performance. This effect appears most striking for these two rates (factor omission, dash-dot grey line and indicator commission, black solid line) in the six factor data (lower right plot).

As will be seen in Figures 18-25, the variances of values for the main effects were not necessarily equal across groupings, so one-way nonparametric ANOVAs (Kruskal-Wallis) were carried out comparing the three main error rates (factor and indicator omission, indicator commission) across the levels of the condition variables. These results are given in Table 4.

Table 4. Main effects for conditions on errors from CFA

	Omission errors- factors	Commission errors- factors	Omission errors- indicators	Commission errors- indicators
Sample size (100, 300, 500)	χ^2_2= 18.36 ***	χ^2_2= 0.39 p=0.8	χ^2_2=0.36 p=0.8	χ^2_2=1.42 p=0.5
Number of factors (0, 1, 4, 6)	χ^2_2= 15.65 **	χ^2_2= 10.79 **	χ^2_2= 35.24 ***	χ^2_2= 90.10 ***
Strength of relations (.8, .6, .4)	χ^2_2=84.39 ***	χ^2_2=2.09 p=0.3	χ^2_2= 0.84 p=0.66	χ^2_2=5.44 p=0.066
Saturation (pure, share 1, share 2)	χ^2_2=5.34 p=.069	χ^2_2=6.47 *	χ^2_2= 124.78 ***	χ^2_2=56.93 ***
Weak factor (yes, no)	χ^2_1= 29.35 ***	χ^2_1= 17.64 ***	χ^2_1=0.01 p=0.9	χ^2_1= 0.39 p=0.53
Structural model (independent, correlated factors)	χ^2_1= 4.98 *	χ^2_1= 0.11 p=.7	χ^2_1=0.13 p=0.7	χ^2_1=15.26 ***
* p<0.05; ** p<0.01; *** p<0.0001				

The results in Table 4 suggest that, when collapsing across the other conditions –including number of factors in the population all conditions, except for saturation, affected CFA's performance at dimensional discovery in terms of omitting factors. Factor commission errors (finding extra factors) were affected by the number of factors, saturation, and the presence of a weak factor. Indicator omission rates were affected by the number of factors and the degree to which factors were pure, while indicator commission rates were significantly affected by the number of factors, independence, and purity of factors.

4.4.2 CFA performance by conditions

Figure 18 presents CFA performance at dimension discovery (factor omission errors) for the four and six factor data across loadings and whether factors were orthogonal. Within each panel, the performance under the three purity conditions (share 0, open diamond; 2 latents share 1, star; 2 latents share 2, filled circle).

As noted, these conditions were not part of the zero and one factor results. There are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and presence of a weak factor.

From the slopes of the lines in each plot shown in Figure 18, it can be seen that factor omission error rates decrease (accuracy increases) as loadings increase (dashed line); this is true whether factors are orthogonal or correlated and for all purity condition levels. This pattern is much simpler than in the CCA results. However, in the six factor data, improvements in factor omissions is shallower (less change as loadings increase) for correlated factors than for uncorrelated factors. This is more pronounced than in the four factor results.

Figure 18. CFA Factor omission errors by independence/loading conditions, four and six factors

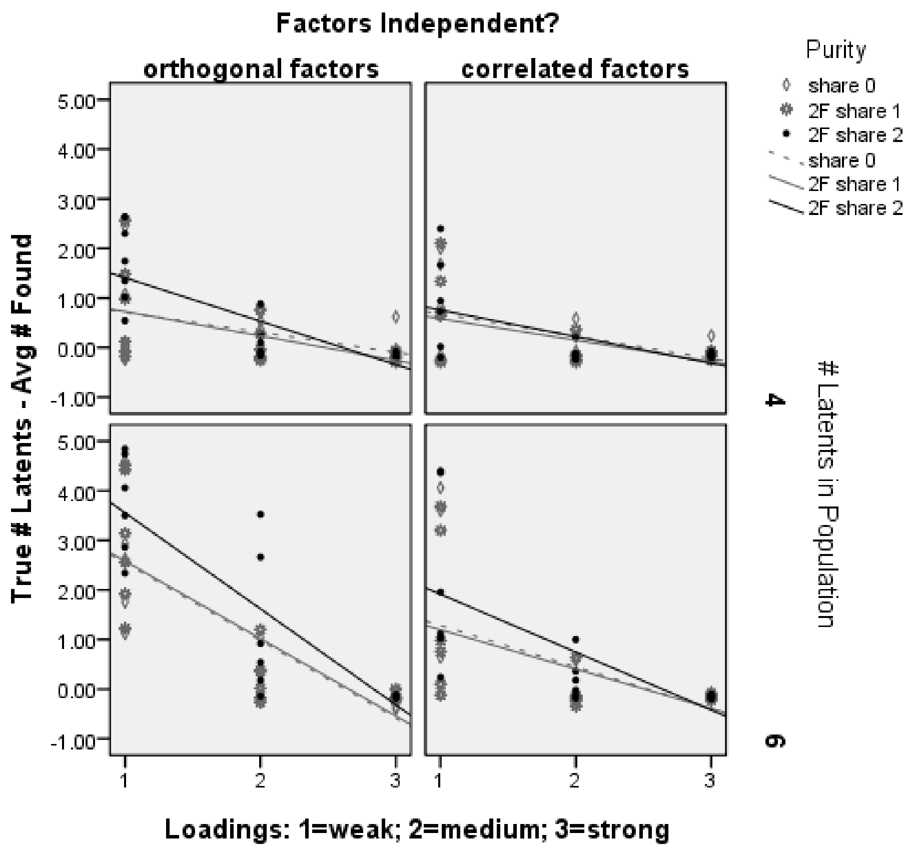


Figure 19 shows the commission error rate for factors across conditions for the four and six factor data. The error rates tend to vary little by purity condition although variability in factor commissions appears to decrease as purity decreases. The effects of correlated factors are not pronounced when factors are pure and when two factors share one or two indicators (solid lines) the factor commission errors tend to respond to loadings (unlike when factors share no indicators, dashed line).

Figure 19. CFA Extra factors found per solution across independence/loading conditions for four and six factors.

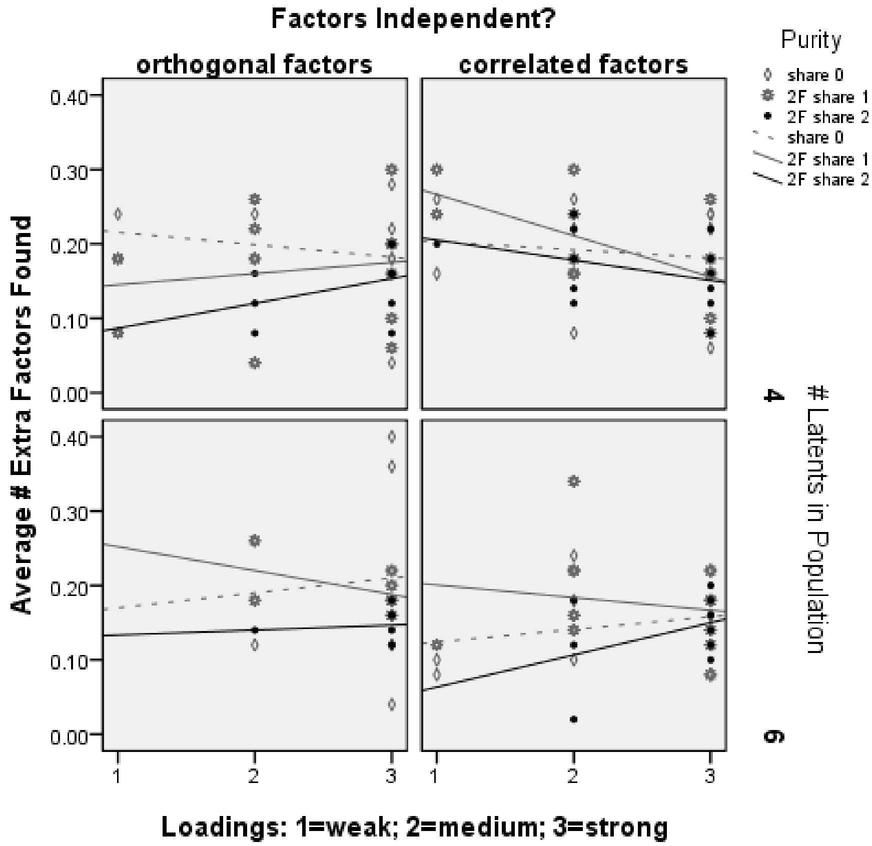
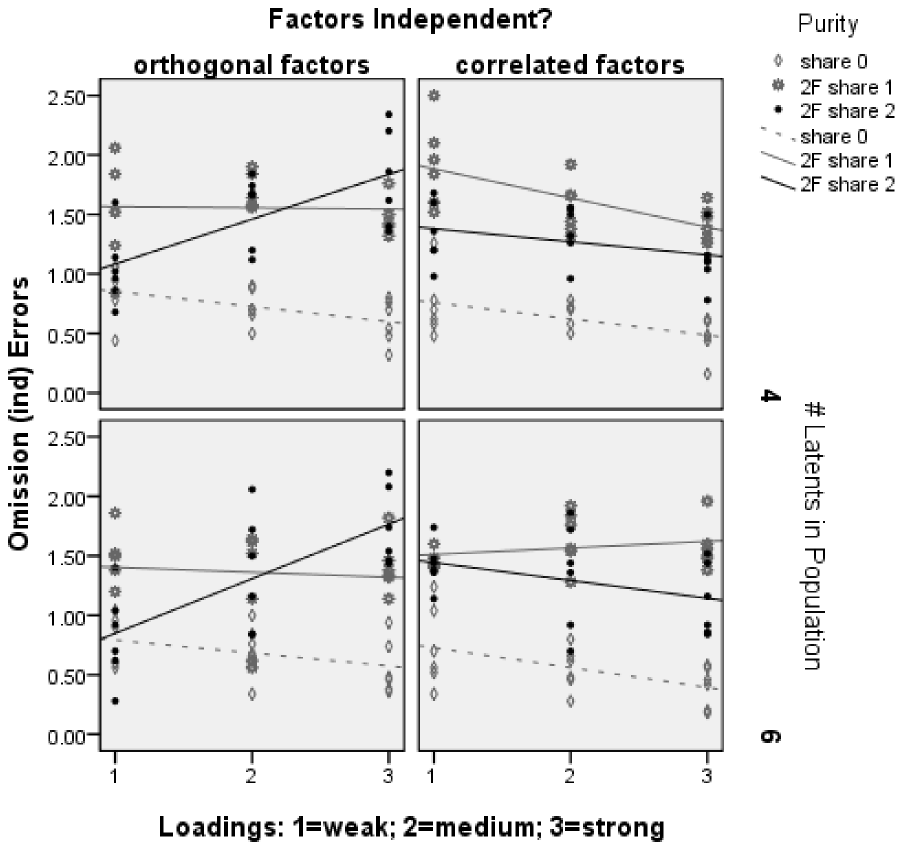


Figure 20 shows the omission rates for indicators. These errors appear unaffected by correlated factors when factors are pure (grey dashed lines) and when two factors share one indicator (grey solid line). When two factors share two indicators, error rates that tended to increase as loadings increase instead were flat, or slightly decreased with increasing loadings. Omitted indicator errors were more common when factors were not pure, although when two indicators were shared, loadings were positively associated with more omitted indicators when factors were orthogonal.

Figure 20. CFA Omitted indicators per solution across conditions for four and six factors.



Factor commission errors for indicators tended to decrease (i.e., CFA performance improved) with strength of loading for both four and six factor solutions. Only the six factor results reflect a sensitivity to shared indicators in terms of indicator omission error rates.

The errors of commission for indicators, where the CFA solution placed an indicator away from its sentinel or either with or without its sentinel but in the minority on a factor, are shown in Figure 21. The commission error rates were far greater than for the CCA results.

When factors were correlated, fewer factors were found; thus lower commission error rates for the correlated factors conditions (right plots) is expected. Like omission errors, sharing one or two indicators did not alter error rates given that the factors were impure (solid lines are very close in all conditions). Unlike the omission errors, however, only commission errors for six orthogonal factors (lower left plot) decreased with loading strength; commission errors were essentially flat in all other contexts.

Figure 21. CFA committed indicators per solution across conditions for four and six factors.

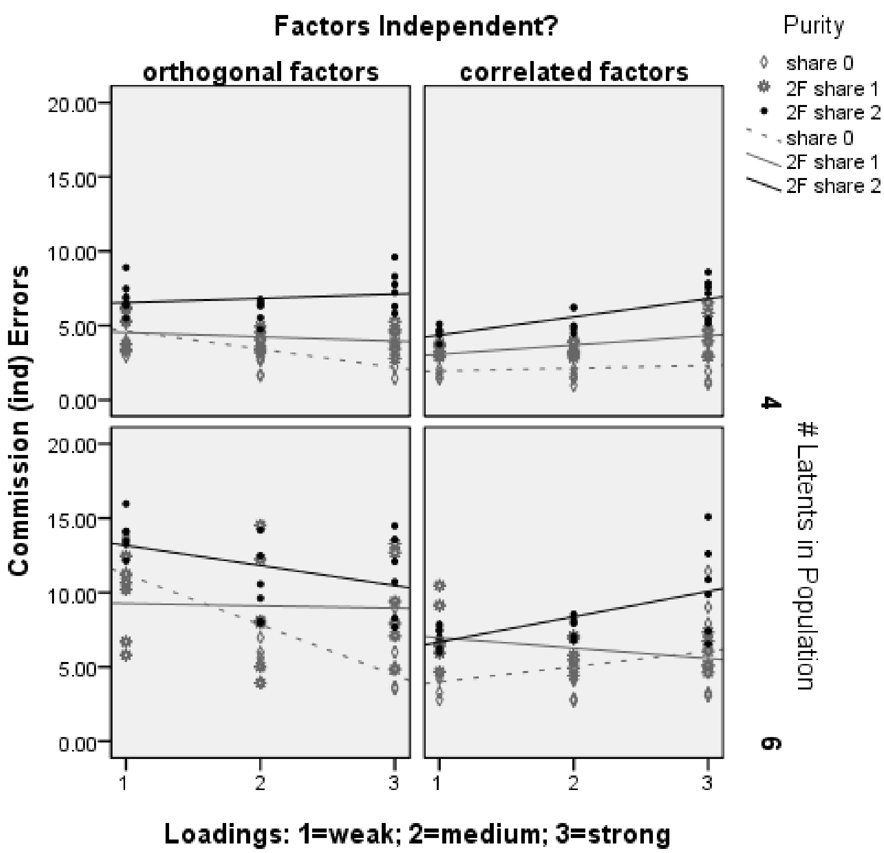


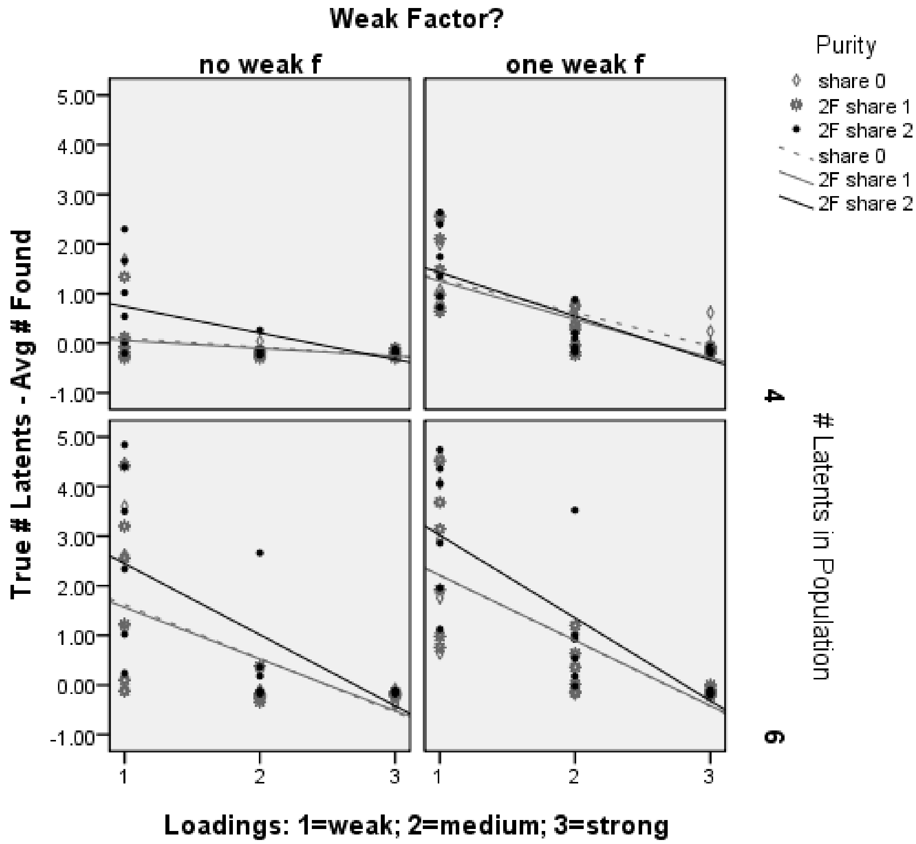
Figure 21 shows the commission error rate for indicators across conditions for the four and six factor data. The indicator commission rates tended to decrease as loading strength increases for both four and six factor data with pure factors when factors are uncorrelated, and the relationship to loadings tends to be stronger in the six, relative to the four, factor data. When any sharing is present (solid lines), commission errors appear to be affected slightly by strengths of loadings in the presence of correlations among factors. Commission errors for indicators are lower when factors are correlated for all cases.

In Figures 22-24, the observations for each loading in each panel represent the six (2x3) combinations of sample sizes and whether factors were orthogonal.

Figure 22 presents CFA performance at dimension discovery (factor omission errors) for the four and six factor data across loadings and the presence of a weak factor. Within each panel, the performance under the three purity conditions. As noted, these conditions were not part of the zero and one factor results. There are six observations for each loading in each panel, representing the six (2x3) combinations of sample sizes and whether factors were orthogonal.

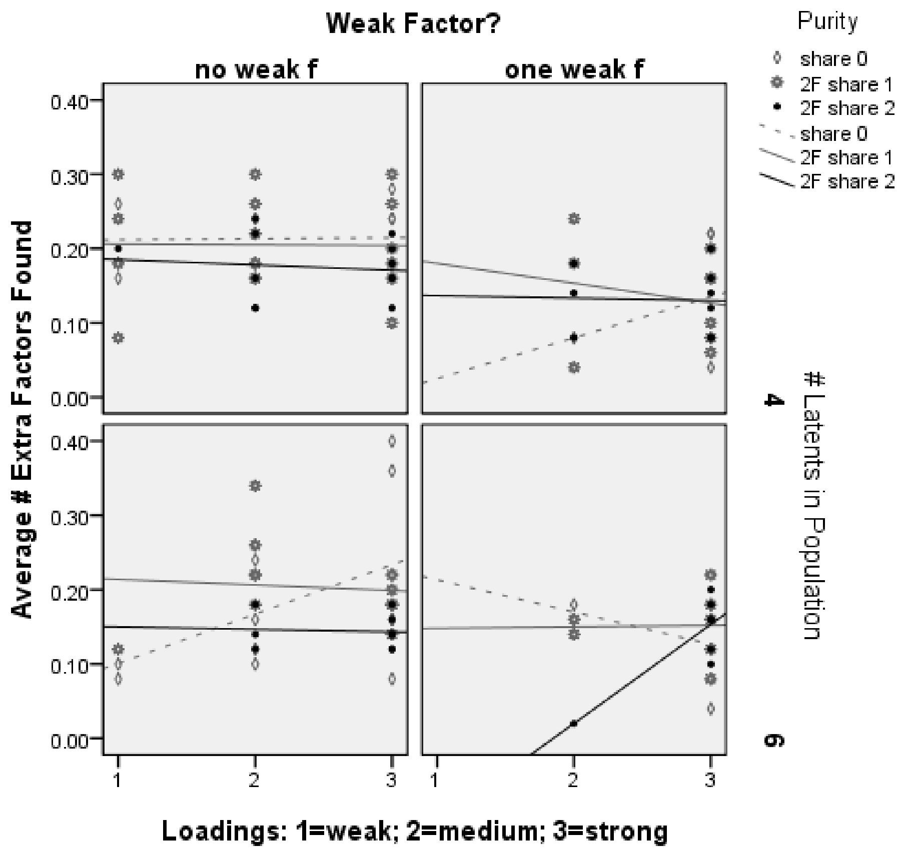
When no weak factor is present, there are small differences in the levels of factor omission errors depending on whether factors are sharing indicators, and in six factor data there is a stronger effect of factor loadings on this error rate. The patterns are similar in the presence of a weak factor, suggesting that its presence did not dramatically affect CFA performance in terms of factor omission.

Figure 22. CFA dimension discovery by loadings/weak factor conditions, four and six factors.



In Figure 23, loadings appear to affect factor commission rates only for six factor data with pure factors when there is no weak factor (grey dashed line, bottom left plot) or when two factors are sharing two indicators when there is a weak factor present (black line, bottom right plot).

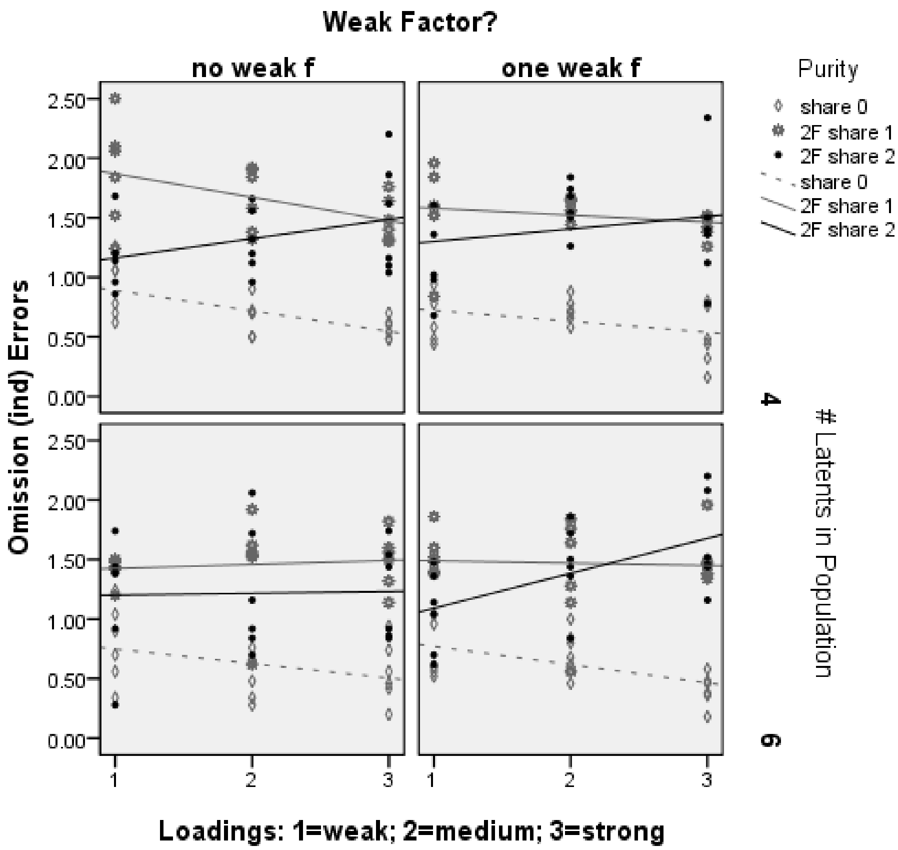
Figure 23. CFA Factor commissions per solution across loadings/weak factor conditions for four and six factors.



In the presence of a weak factor when factors are pure, factor commission errors increase with increasing loading for the four factor data only.

Figure 24 shows that, whether there was a weak factor or not, factors sharing indicators tended to increase the omission of indicators in the CFA solutions. The effects of a weak factor were only observed on this type of error when one (grey solid line) or two (black solid line) indicators were shared by two factors in the population.

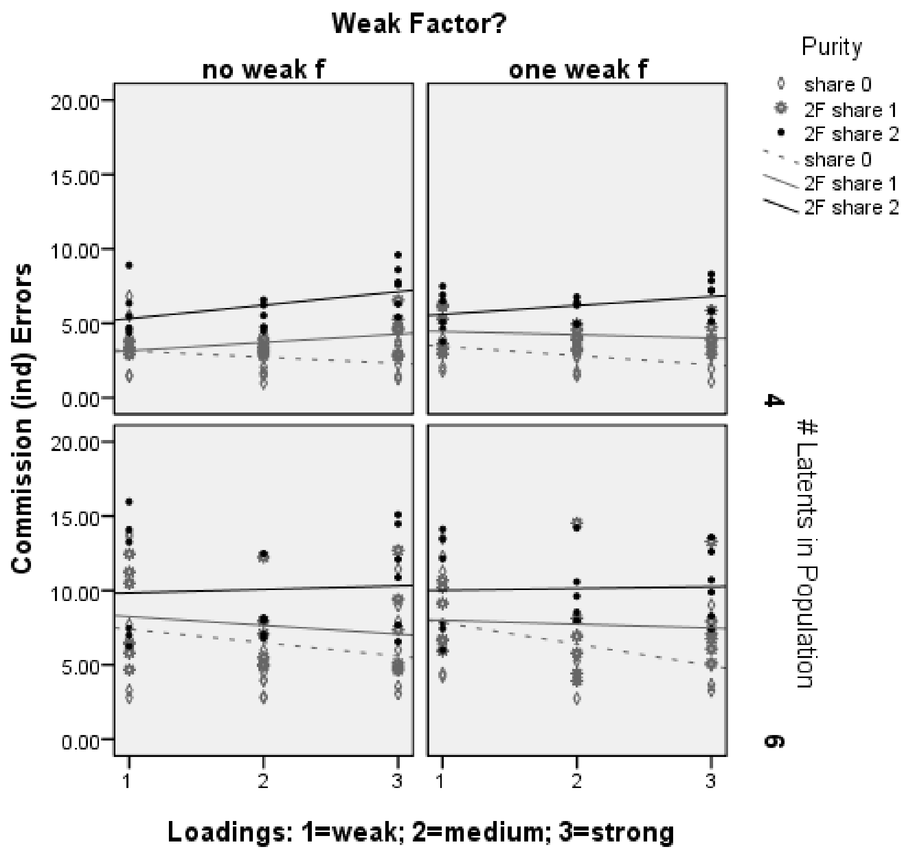
Figure 24. CFA Omitted indicators per solution across loadings/weak factor conditions for four and six factors.



The errors of commission for indicators, where the CCA solution placed an indicator away from its sentinel or either with or without its sentinel but in the minority on a factor, are shown in Figure 25. Commission of indicator error rates were the highest of the four types observed in the CFA results.

Like indicator omission errors, sharing one or two indicators did not alter commission of indicator error rates in four factor data given that the factors were impure (solid lines are very close in all conditions). Indicator commission errors were essentially flat across loadings and did not appear affected dramatically by the presence of a weak factor. In the four factor data, sharing increased indicator omission errors relative to pure factors, and this was attenuated slightly in the presence of a weak factor.

Figure 25. CCA committed indicators per solution across loading/weak factor conditions for four and six factors.



The six factor commission of indicators errors was very similar with and without the weak factor, with a slight decrease in errors as loadings increased in the presence of a weak factor.

Like the CCA results, these CFA results are focused on accuracy since with the exception of the perfect performance (of confirmatory factor analysis) on the one factor data, consistency was very low. The omission of factor errors were significantly impacted by all conditions except saturation (purity) of factors, while commission of factor errors were significantly impacted by the number of factors (mainly due to an extreme commission error rate for the zero factor data analyzed by confirmatory factor analysis), purity of factors, and the presence of a weak factor. For indicators, the commission error rates were much higher than the omission errors; the effects of the number of factors and the saturation conditions were significant for both omission and commission of indicator errors, while correlations of factors impacted only the indicator commission error rate.

4.5 Comparison of Results by Method

4.5.1 CCA vs. CFA at dimensional discovery

The results (error rates) for the two methods were compared by a series of two-way ANOVAs so that effect sizes could be computed as planned (in Chapter 3). The analyses were also carried out via nonparametric methods (data not shown) and inferences were not different; parametric analysis results are given for consistency with the effect size reporting (since η^2 values are not available from nonparametric analyses). Significant interactions for method with most of the conditions for at least one of the four error rates given in the table render the main effects less interesting, but all main effects are presented in Table 5A for completeness.

The results in Table 5A suggest that, when collapsing across the other conditions –including number of factors in the population, CCA made significantly fewer errors of commission and CFA made significantly fewer errors of omission than the other method, respectively. This is the interpretation of every significant main effect in Table 5A and is consistent with the conclusions from earlier comparisons of CCA and CFA (Silva, 2005).

Table 5A. Statistics for main effects and effect sizes on errors by method for conditions

	Omission errors- factors	Commission errors- factors	Omission errors- indicators	Commission errors- indicators
Sample size (100, 300, 500)	F(2, 450)=18.47 *** η^2=0.076	F(2, 450)=5.04 ** η^2=0.022	F(2, 450)=5.32 ** η^2=0.023	F(2, 450)=0.84 ns
Number of factors (0, 1, 4, 6)	F(3, 448)=14.69 *** η^2=0.090	F(3, 448)=45.46 *** η^2=0.233	F(3,448)=30.29 *** η^2=0.169	F(3, 448)=54.94 *** η^2=0.269
Strength of relations (.8, .6, .4)	F(2,444)=52.92 *** η^2=0.192	F(2, 450)=39.74 *** η^2=0.152	F(2,444)=13.25 *** η^2=0.0256	F(2, 444)=8.4 *** η^2 0.038
Saturation (pure, share 1, share 2)	F(2,426)=52.92 *** η^2=0.192	F(2, 426)=2.56 ns	F(2,426)=28.75 *** η^2=0.119	F(2, 426)=24.12 *** η^2=0.102
Weak factor (yes, no)	F(1,428)=26.98 *** η^2=0.059	F(1,428)=41.73 *** η^2=0.089	F(1,428)=3.63 ns	F(1,428)=.96 ns
Structural model (independent, correlated factors)	F(1,428)=4.25 ** η^2=0.01	F(1,428)=2.87 ns	F(1,428)=4.53 * η^2=0.010	F(1,428)=29.05 *** η^2=0.064
* p<0.05; ** p<0.01; *** p<0.0001				

The results in Table 5A support conclusions from the method-specific results that, when collapsing across the other conditions –including number of factors in the population, CCA made significantly fewer errors of commission and CFA made significantly fewer errors of omission than the other method, respectively. This is the interpretation of every significant main effect in Table 5A and these results are consistent with Silva's (2005) findings.

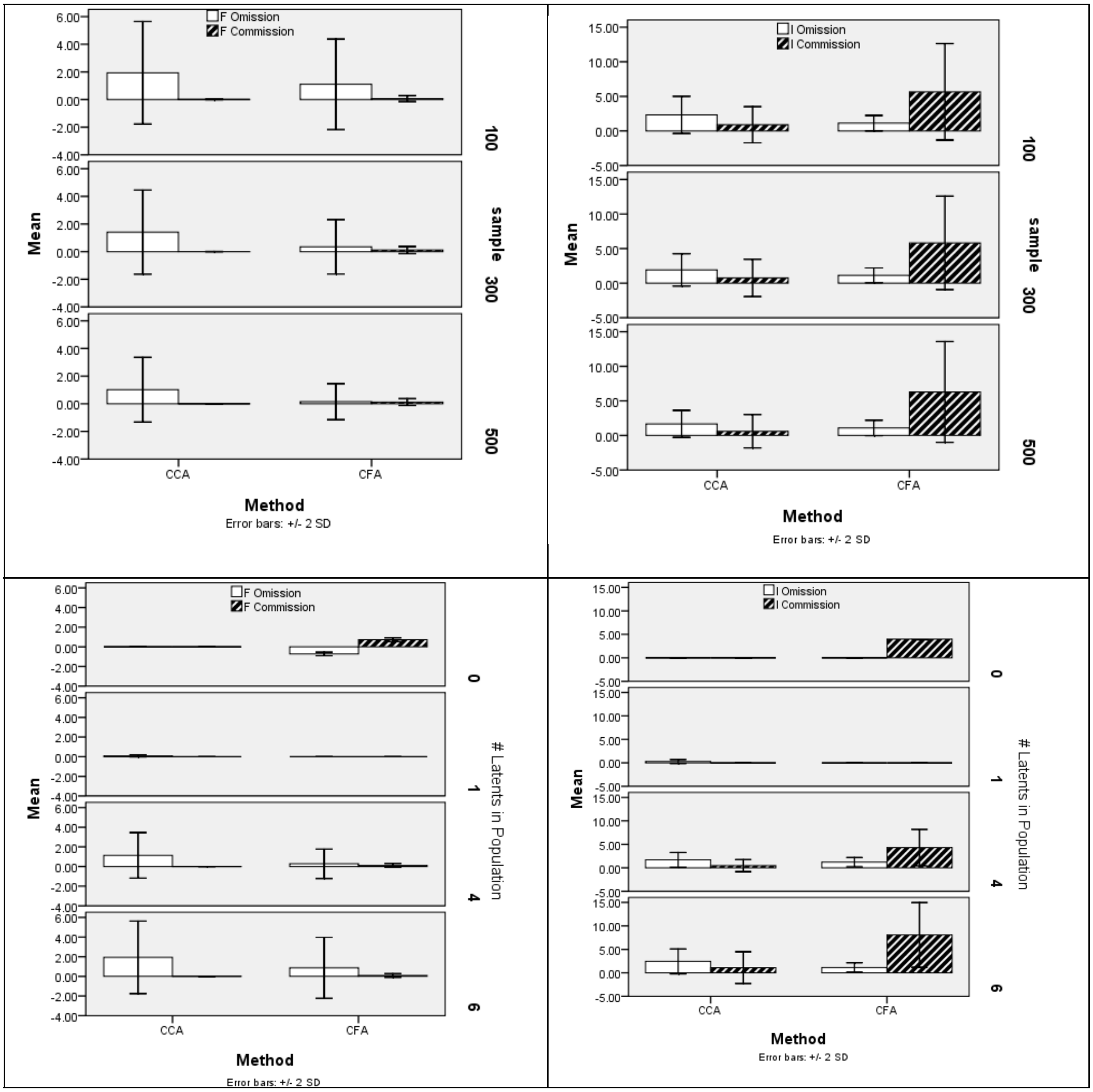
Table 5B. Interaction effects between Method and Conditions for four error rates.

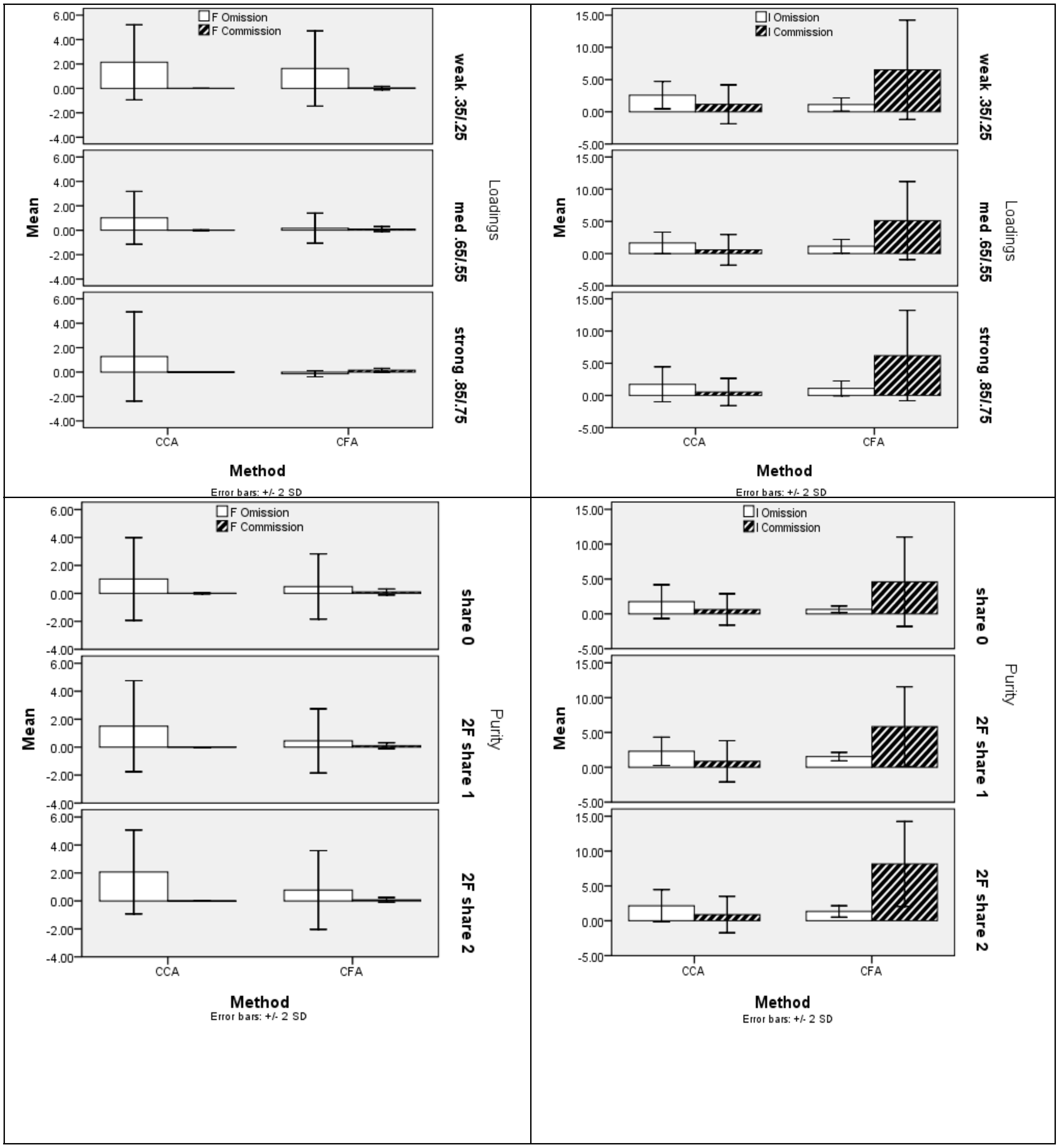
	Omission errors- factors	Commission errors- factors	Omission errors- indicators	Commission errors- indicators
Method x Sample size	ns	F(2, 450)=5.24 ** η^2=023	F(2,450)=4.32 * η^2=.019	ns
Method x Number of factors	ns	F(3, 448)=45.82 *** η^2=.235	F(3,448)=9.71 *** η^2=.061	F(3, 448) = 28.93 *** η^2=0.162
Method x Strength of loadings	F(2, 444)=4.84 ** η^2=0.021	F(2, 450)=5.24 *** η^2=0.146	F(2,444)=12.92 *** η^2=.055	ns
Method x Saturation	F(2, 426)=4.84 ** η^2=0.021	ns	ns	F(2,426)=18.73 *** η^2 =0.081
Method x Weak factor	ns	F(1,428)=41.34, *** η^2=0.088	ns	ns
Method x Structural model	F(1,428)=26.98 *** η^2=.059	F(1,428)=4.68, * η^2=0.011	F(2, 426)=5.81 * η^2=0.013	F(1,428)= 10.347 ** η^2=0.024
* p<0.05; ** p≤0.01; *** p<0.0001				

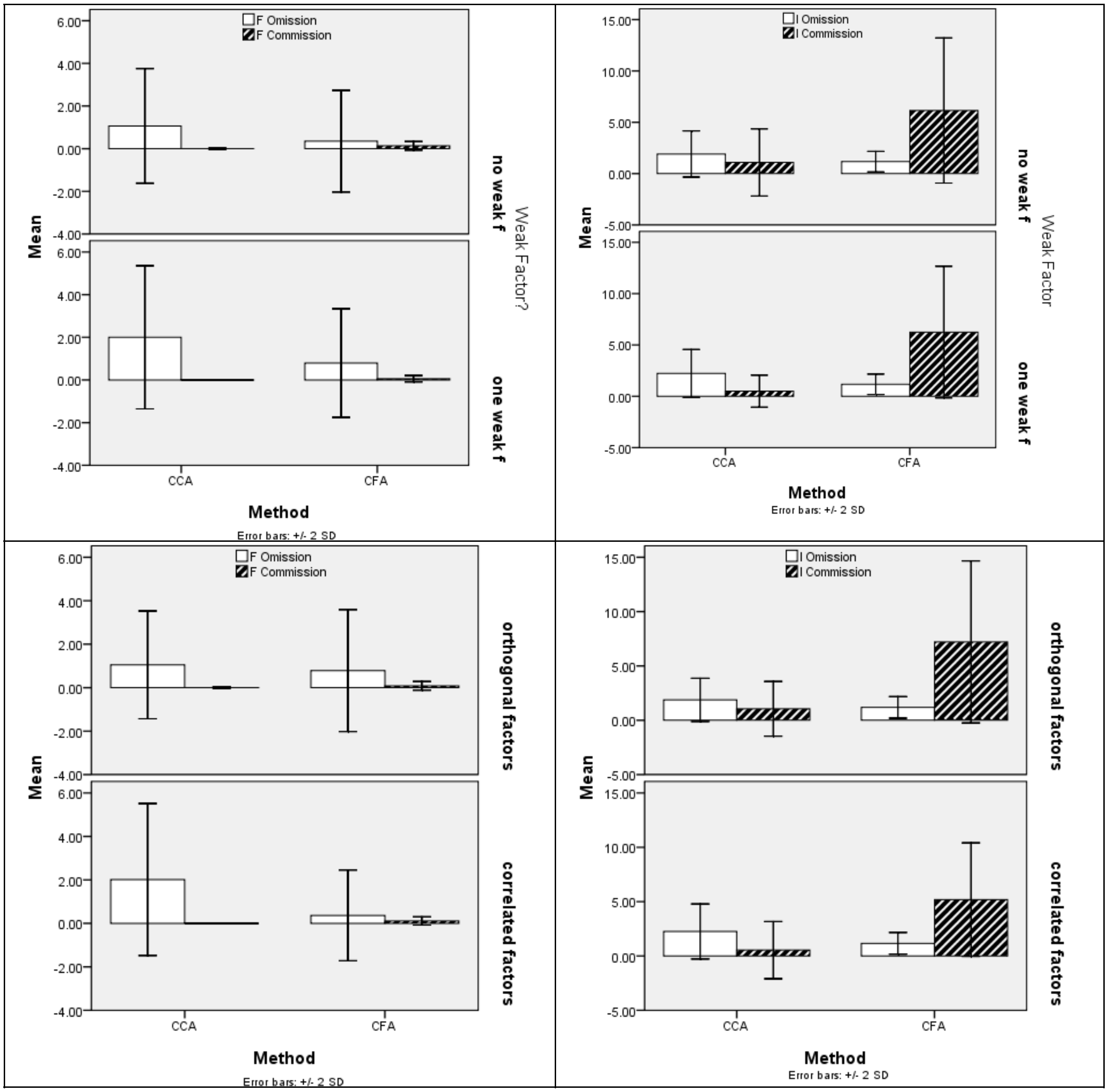
Table 5B shows that the correlations-between-factors condition interacted with method in terms of all error rates and strength of loading interacts with method for all but commission of indicator errors. The effects of these conditions, and their interactions with method in terms of their impacts on the error rates, can be generally characterized as varying in strength but not direction; the main effect of method in every analysis was significant and the error rates were always lower for commissions by CCA and higher for omissions by CCA relative to CFA. Thus, the main effects of these conditions and their interactions with the method of estimating the dimensions in a data set all lead to the same conclusion: CCA tends to extract fewer, and assign fewer indicators while CFA will tend to support more, and assign more indicators, in any factor model. Conditions that improved the performance of one method tended to worsen the performance of the other, resulting in significant interactions for five of the six conditions with the method in terms of commission errors for factors and for four conditions in omission errors for indicators. Purity and correlations between factors led to significant interactions with method for omission errors for factors and commission errors for indicators, the two types of errors most particular to each method (commission: CFA, omission: CCA).

In Figure 26, the factor- (left column) and indicator- (right column) based error rates are presented for each method, stratified by each condition separately in bar graphs showing the mean value, with error bars representing two standard deviations around the mean. For all rows except the first two (where stratification is by the sample size (first row) and number of latent variables in the population, second row), only four and six factor data are represented. Within each plot, the error rates are collapsed across the other conditions.

Figure 26. Effects of conditions on errors (factors, indicators) by method.




F Omission
F Commission
I Omission
I Commission
Mean
Loadings
weak .35/.25
med .65/.55
strong .85/.75
Purity
share 0
2F share 1
2F share 2
CCA
CFA
Method
Error bars: +/- 2 SD



Overall, the plots in Figure 26 reflect the greater tendencies of CCA to make omission type errors and of CFA to make commission errors. The plots also reflect the method by condition interactions identified in Table 5B. Sample size tended to decrease factor omission by both methods but decreased factor commissions by CCA while increasing these in CFA. Increasing sample size tended to decrease the indicator omissions by CCA while having little or no impact on CFA indicator errors.

The plots stratified by number of latent variables in the population show the increase in errors with increasing numbers of latents, although for factor commission and both indicator type errors, the methods were significantly different in terms of the response to this condition (i.e., there was a significant method by number of latents interaction). The plots by loadings show the tendency for errors to decrease as loadings increased, although the indicator omission rates by CFA when loadings are medium or strong are very similar and method interacted significantly with loadings for the factor type errors. A significant interaction between method and purity is shown in the commission errors for indicator increasing with decreasing purity in CFA results and not varying in CCA results; a similar interaction is observed in the omission of factors increasing with impurity in CCA results and tending not to change in CFA results.

The presence of a weak factor interacted with method only in factor commission errors, where adding the weak factor decreased factor commissions by CFA but did not affect CCA results. All error types interacted significantly with method when factors were either orthogonal or correlated, with unchanging or decreasing indicator omission errors in CFA results and increasing indicator omissions in CCA. The indicator commission errors by CCA changed little when factors were correlated relative to uncorrelated, while adding correlations between factors reduced CFA indicator commission errors more substantially. Factor omission errors by CCA increased while factor commission errors decreased across the factor independence condition; the opposite patterns were observed for CFA results, leading to these significant method-by-independence interactions.

4.6 Summary of Results

4.6.1 Did CCA and CFA perform comparably and did conditions significantly impact either method?

In Chapter 3 the purpose and plan for the study was laid out in a table (Table 2). Given the foregoing results, the answers can be filled in.

Table 6. Answers to questions of substantive interest for study

Design characteristic	Question
	Main effects
Method	All error rates were significantly different across methods: omission type errors were significantly greater for CCA and commission type errors were significantly greater for CFA.
Sample size (100, 300, 500)	Errors tended to be lower as sample size increased over all; this was a significant (expected) trend. Commission errors for indicators were not sensitive to sample size.
Number of factors (0, 1, 4, 6)	The error rates were significantly different across the number of factors in the population. This was not unexpected since more factors brings more opportunities for all types of errors.
Strength of relations (.8, .6, .4)	Errors were significantly different across strengths of relations, with lower error rates when relations were stronger. (see interaction note)
Saturation (pure, share 1, share 2)	Purity of factors was significantly associated with greater errors of omission and commission for factors only.
Weak factors (yes, no)	The presence of a weak factor was significantly associated with greater errors of omission and commission for factors only.
Structural model (independent, correlated factors)	Correlations (0.50) among factors was significantly associated with omission and not commission of factor errors (see interaction note).

Interactions	
Method x Sample size	A significant method by sample size interaction was observed for commission errors for factors and omission errors for indicators.
Method x Number of factors	This significant interaction was driven in part by the increasing number of opportunities for making errors as factors increased and the different types of errors each method made when given the opportunity, and in part by the disproportionately high errors on zero factor data by confirmatory factor analysis.
Method x Strength of loadings	Strength of loadings is a key feature for CFA estimation and is not as important in CCA estimation since CCA uses proportionality in covariances rather than size. This interaction was predicted.
Method x Saturation	The CCA algorithm is designed to exclude indicators (and subsequently, their factors) when factors are not pure while CFA utilizes this information in its estimation. This interaction was predicted.
Method x Weak factor	Only commission of factor errors exhibited this significant interaction, since CFA tended to over-extract factors while CCA missed them (and any factor with a sub-threshold correlation) in the presence of a weak factor.
Method x Structural model	The presence of correlations among factors affected all error types by increasing one type and decreasing the other for the two methods.

Table 6 summarizes the results presented in this chapter. Discussion of their interpretation, and issues pertaining to the estimation and interpretation of these results, as well as next steps for the research on this topic, are provided in Chapter 5.

Chapter 5: Discussion

The discussion of this study focuses on three domains: the results themselves, the interpretability of the results and whether the methods are or are not comparable, and implications of these results for next-step research. In general, the criteria for interpreting the relative performances of these methods were very specific- either to the method, to the desired comparisons across such divergent methods, or to the situations that the simulation conditions might reflect. The discussion section is therefore also a consideration of the limitations of the study.

5.1 Results, and their interpretation, depend on the method

5.1.1 Relative performances by the two methods

Overall, CCA (BuildPureClusters) was perfect (accurate and consistent) in estimating the true number of factors in the zero factor case; CFA (confirmatory analysis, the proxy for common factor analysis in this study) was inaccurate but very consistent in the zero case. CFA was perfect in the one factor data, across sample sizes and loading strengths; CCA was affected by loading strengths in the one factor data. Overall, CFA committed significantly fewer omission errors, and significantly more commission errors, for factors than CCA did; similar significant differences were obtained for the omission and commission errors for indicators. Neither method attained the invariance criterion of 95% of the 50 trials returning the same model beyond the one factor data in both (except the N=100 condition for CCA) and the zero factor data in CCA. Thus, the answer to the main research question of this study (to determine if CCA was better than, or equal to, CFA) is that the methods are equally good (or bad) and selecting one method will depend on the investigator's purpose and the implications for errors of omission and commission.

Under-extraction is generally considered worse, and more consistently damaging, result of mis-specification of the true dimensionality of the variable/score set, than over extraction (O'Connor, 2000; Velicer, Eaton, & Fava, 2000). However, the difficulties of under-extraction these authors describe come mainly from the implications for loadings, not for understanding the latent structure of the data, *per se*, although the example shown in Figure 1 (Chapter 1) reflects the potential impact of under-extraction on future science.

On the other hand, the algorithms in TETRAD are designed specifically not to conclude that a common cause exists if insufficient evidence for its existence is found. Thus, on the surface, TETRAD was designed to under-extract if it is used to estimate the dimensionality of a set of variables or scores. Based on the perspectives of its developers, it is designed to provide evidence for causal relationships and when it does not find this evidence, whether that is due to insufficiency or sufficient evidence of no relationship, it returns "no evidence". This is an important – critical – distinction between the omission errors (for factors or indicators) from CFA and CCA. The comparability of the methods is considered in the next section.

Another important feature, which was common to both CCA and CFA, was the definitions of errors (commission, omission) used in this study. The intention was to simulate data – and make decisions about loadings – that corresponds to what practitioners usually do. All solutions were considered, and unlike other modelers who have used CCA (via TETRAD), all errors from all solutions were counted. This resulted in a very large number of errors for both methods, unlike error rates in previous studies. However, like in other studies, this analysis found that CCA tends to make more omission type errors while CFA tends to make more commission type errors. The interpretation of non-convergent confirmatory one-factor models as supporting a zero-factor solution led to 10%-40% correct CFA results; without this interpretation, the CFA results would have been none correct for zero factor samples. As described in Chapter 2, CFA (common or confirmatory FA) cannot work if m=0; but giving 10-40% correct to CFA in the zero factor condition does not make it comparable to CCA, which was 100% correct in this condition, so this choice did not affect the conclusions a great deal.

The results of this study for the other conditions were similarly not necessarily dependent on the method of error computation – the errors were defined in a consistent manner across the methods and conditions. Importantly, all opportunities for error counting were used (unlike Silva, 2005 where "improper" solutions were not included in error computations). It is unlikely that different rules for error definition would lead to different results, especially given the consistency of these results with Silva's 2005 outcomes, in spite of very different scoring rules in this and the 2005 study.

Choosing one method over the other will depend on the investigator's willingness to either over- or under-extract. Although not examined in this study, TETRAD can uncover structural relations between latent variables, which CFA cannot do (because it assumes a first order model). So, in addition to considering the implications of under- or over-extracting of factors and the commission or omission of indicator assignment, investigators should consider their ultimate modeling goals. If structural relations are to be modeled (or uncovered) among unidimensional measurement models, TETRAD will work well- given the correlations are strong enough and the sample sizes are large (300 or larger).

5.1.2 Dimension estimation performance

Searching for a minimum number of factors is *inconsistent* with a search for the *correct model* (as very carefully outlined by Hayduk & Glaser, 2000), and a model that fits the data “adequately” is not necessarily the correct model (Bartholomew, 2007, p. 13) – CCA as implemented in this study reflects these perspectives. An interesting corollary to this perspective is proposed by Herting and Costner (2000), who argue that determining the proper number of factors, while originating from “the explanatory hypotheses that motivate the model”, is simply a method to “improve the fit between model and data” (p. 93) – CFA as implemented in this study reflects these perspectives.

Estimating the “proper” number of factors using correlation constraint analysis is not related to the fit of a model to the data; that is only an appropriate characterization for analysis techniques such as common and confirmatory factor analysis, or methods based on their outcomes (such as log ratio or fit statistics using the model χ^2). The results of this study suggest that the two methods are biased towards under- (TETRAD) and over- (AIC/ECVI from CEFA) extraction type errors, so at a minimum it might be advisable to run both types of analysis to determine a likely level of dimensionality for a given set of data. If the CEFA and TETRAD results diverge, investigators should consider confirmatory factor analysis to test specific hypotheses about the underlying structure and relationships.

CCA performance was sensitive to each condition, including the presence of a weak factor. As was noted in the Results chapter, TETRAD requires correlations among observed variables to be 0.30 (Spirtes, et al. 2000); higher errors were observed for the weak loadings condition and with a weak factor having 50% smaller loadings than the other factors, the presence of a weak factor would have been difficult to detect in both the weak and medium loadings conditions. It is also possible that CCA performance was actually more susceptible to experimental conditions when there were more factors and stronger loadings. Some errors made in the strong-loading condition might be most easily attributable to the conditions, whereas at least some of the errors in lower-loadings conditions must be attributable in part to the threshold of TETRAD's sensitivity. These divergent attributions cannot be resolved with the data or results presented in Chapter 4. However, a future study could be devised specifically to test this hypothesis (of error attribution).

Importantly, any study of TETRAD will require, as these simulations did, four indicators per latent variable, unidimensional latent variables, and sufficient population correlation coefficients. The performance capabilities of CCA (implemented via TETRAD) without these key requirements cannot be determined since the software cannot operate without the required features in the data. Any simulation study assumes some baseline level of conformity with assumptions; a simulation study with vanishing tetrads assumes at least four indicators and with BuildPureClusters, it assumes unidimensional factors.

The results suggest that correlation constraint analysis is better than confirmatory factor analysis at identifying the *minimum* number of common, unidimensional factors when there is no common factor in the population; as articulated in Chapter 3, common factor analysis does not work when there are zero factors, so it clearly cannot be used to determine whether zero or one is the true minimum number of factors. Given the interpretation of a failure of the confirmatory one-factor analysis to converge as a correct representation of zero factors under this method, the results in Chapter 4 show that data from a zero factor population were often (over 60% of the time) fit by a one factor confirmatory solution. By contrast, correlation constraint analysis (TETRAD's BuildPureClusters algorithm) returns a "zero factor" solution when there is evidence of no common latent factor, when there is evidence of a single common factor that is *not* latent, when there is evidence that indicators have >1 cause, and when there is insufficient evidence of one causal common factor. As such, investigators should use it for unidimensional measurement models and in conjunction with confirmatory factor analysis in order to determine why a zero factor model was returned.

5.1.3 Success of the sentinels for identifying omission errors for indicators

As outlined in Chapter 3, counting omission and commission of indicator errors required the determination of which factor in the population a set of indicators grouped together in a solution was representing. In this study the first indicator on each factor in the population was treated as a sentinel, so that when that indicator was observed, the other indicators that should have loaded with it would be known and errors could be classified. The sentinels were useful in decisionmaking for error classification and counting, and there was no characteristic of the methods that would give rise differentially to errors of one sort or the other, given the use of sentinels in decisionmaking. In some solutions (as described in Appendix 2) there was no sentinel assigned to a factor; in other solutions the factor was only comprised of sentinels. Designating sentinels was a useful tool that *facilitated* the analyses and permitted similar information to be used from CCA and CFA solutions for classifying errors – but a classify-factor-according-to-majority-indicators rule was also important. However, their use suggests that these results might not generalize to cases where all items load equally on their common factor; since these results are consistent with earlier simulation results and because the true structure is unknown in real data, the implications for using sentinels is unclear, and could be quite minimal.

In his simulation study, Silva (2005; Silva, et al. 2006) used a simple majority rule to determine which factor a set of indicators represented for the CCA method results, but used a different rule for the factor analysis results: the absolute values of the loadings of indicators in a solution (grouped according to respective true latents) were summed and whichever true latent the largest absolute value of the sum pertained to was designated the latent that those indicators represented; the other indicators were classified as commissions and missing (from the true latent) indicators were identifiable. Unlike Silva's 2005 study, this study specifically sought the same type of information from solutions from each method, such that any factor loading from a CFA solution was retained (=1) if it was statistically significantly different from zero (loading/SE(loading)>1.96), otherwise it was excluded from consideration. The results could have been different if other choices had been made for error counting – but the conclusions (CCA tends to under-extract and omit indicators/CFA tends to over-extract and commit indicators) were probably not dependent on these error counting rules since these results were similar to Silva's (2005).

5.1.4 Zero factor condition

The study included a zero (uncorrelated) condition, although zero factors are impossible for CFA to estimate. To circumvent this problem, confirmatory analyses, fitting a one-factor model, were used to determine whether a one or a zero factor solution would be better fit to the zero and the one-factor samples. As described earlier, maximum likelihood estimation was used in order to obtain standard errors for the pathweights in these solutions. This was deemed acceptable, given the demonstration by Briggs and MacCallum (2003) that maximum likelihood is less accurate that ordinary or unweighted least squares only when weak factors are present (i.e., not for these one-factor models). AIC values for models were used to select the solution with the better-fitting number of factors, i.e., to determine the number of found factors in these samples – just as AIC was used to determine which CEFA model (i.e., number of factors) was the best solution for the 4-6 factor samples.

The results from the confirmatory analyses are attributed to the common factor analysis method, which is reasonable because there was no other way to test these data in a model-implied-covariance estimation procedure. Also, all of the not-CCA (i.e., CEFA + Mplus) solutions involved computation of AIC to determine the number of factors in the solution. The estimation of the fit of the model to the data was not achieved in the same manner, but the similarities outweigh the differences.

The error counting rules were applied to solutions where AIC supported a one factor model over the zero factor (null) model, but where the estimated loadings of the indicators were not significantly different from zero. In all but two cases where AIC supported a one-factor model from a zero-factor sample, no factor loadings were estimated. Thus, the error counting facilitated obtaining results from all solutions.

5.3 Are CCA and CFA Results Comparable?

5.3.1 The results ARE comparable

The introduction (Chapters 1-2) outlined the differences between the estimation and philosophical perspectives behind the two methods considered in this dissertation. The results are comparable in the sense that an investigator who wishes to understand the causal structure of a set of variables should consider all relevant tools. These results do not suggest that one or the other of these methods is not relevant to the problem of estimating causal structure. On the contrary, the arguments leading to the proposal and execution of this work is that an investigator must choose the right tool for the job, and knowing that TETRAD (or CCA) is an option is important for full consideration of all tools, as well as what the "job" really is.

Many reasons for considering the results *not comparable* are articulated in Section 5.3.2. These reasons are focused on quantitative, and not qualitative comparisons. It is clear from these results (Chapter 4) that, if zero or one are plausible numbers of causal latent variables, CFA (or confirmatory factor analysis) are less accurate and less consistent than CCA. Whether the comparison of accuracy and consistency reaches statistical significance in an inference test is immaterial: one method can, and the other method cannot, perform (well) in these contexts, but both methods will ‘try’ to provide some solution. Both methods will appear to work, but only CCA can do what it is supposed to do to return zero or one common factor; CFA will always return (at least) one common factor. Similarly, if an investigator is interested in identifying causal latent variables with simple (pure) structure, common factor analysis cannot perform better than CCA, since CCA is specifically searching for latent causes (vanishing tetrads) and CFA is specifically estimating the correlations between indicators and the pre-specified number of common factors. (This particular feature of CCA was not tested in this dissertation.) Thus, the two methods are comparable because results from both methods can be productively discussed together, and choices made about next steps for analyses on the basis of results from both methods when considered together, particularly in terms of factors and dimensionality. This is true even though the two methods require different features in the data.

5.3.2 The results are NOT comparable

As was articulated in Chapter 4, the two methods (CCA, CFA) led to different challenges in terms of computing the number of errors committed by each one under the diverse experimental conditions this dissertation included. Similarly, these methods require features in the data (e.g., $m>0$ for CFA, at least four indicators and simple structure for CCA). The opportunities for indicator-related errors, and especially the intepretation(s) of omissions, suggest that comparing the results would not be particularly useful. This perspective makes sense in terms of inference testing: carrying out an ANOVA, and estimating effect sizes, when errors from such diverse methods (and when the errors can mean such different things) is not likely to be the most appropriate approach to considering, or estimating, their differences. By extension, if the elements that were analyzed across methods (i.e., errors) were not comparable, then the analysis of those elements will not be interpretable.

Unlike the results from the indicators, the results from the dimension discovery were more comparable. Although the methods estimate the number of factors in a solution in very different ways (see Chapters 2 and 3), the number of factors "found" by the methods are arguably similar. There remains, however, the difference between what an "omission" error represents within each method: as defined in this study, an omitted factor in a CFA solution represents a better fit to the data without that factor. Within CCA an omitted factor either means evidence that the factor does not exist *or* insufficient evidence of its existence. This suggests that the omission errors for factors are also not strictly comparable.

The commission errors for factors and for indicators are highly comparable in the two methods. The difficulty here is only for indicator omission errors: the BPC algorithm in TETRAD for CCA cannot assign one indicator to more than one factor, effectively limiting its opportunities for commission errors, and this is not the case for the CFA results. This suggests a sense of bias, favouring CCA by somewhat artificially limiting the number of commission errors for indicators.

5.4 "Errors" are of different types and entities under the two methods

5.4.1 Comparison, and comparability, of errors for CFA and CCA

As noted in section 4.1.2, it was decided that omission error totals for the solutions from CCA and CFA would include only omissions on the correct (or fewer) number of factors; the first 4 or 6 factors were used and the omission errors occurring in extra factors were ignored throughout. Commission errors for indicators were also computed similarly for CCA and CFA results as the number of commission errors divided by the number of times that factor was found- whether the factor was less than, equal to, or greater than the correct number of factors in the population giving rise to that sample. Other studies involving TETRAD and/or BPC have not included results, i.e., counted errors from, factors that were uncovered if those factors had fewer than 3 indicators assigned to them. In this study, latent variables were not omitted from error counting if they had fewer than 3 indicators (as Silva did, because this is an “improper” solution from TETRAD’s perspective), resulting in the potential for 2-3 additional omission errors per found factor and the potential for more commission errors as well. This is the main reason why these results reflect so much higher error rates than have previously been reported (Silva, 2005; Silva, et al. 2006).

5.5 Recommendations

Herting and Costner (2000) employed a TETRAD analysis to estimate the number of factors in a data set and erroneously describe the outcome: TETRAD output does not recommend that any variables should be eliminated from a measurement model; TETRAD output describes whether there is evidence for a single latent cause for an observed variable, and in the BPC algorithm, if there is sufficient evidence of a second latent cause, then even if the primary cause is the factor of interest, TETRAD will omit that indicator from a ‘pure’ measurement model. A similar failure to understand the results from this software was reported by Yu, et al. (2007). These two reports exemplify how important a study articulating what TETRAD can and cannot do, and how it works, is a potentially important contribution to the social science literature. Further, the interpretation of missing paths in the BPC function of TETRAD is critically dual- indicators could be omitted in a solution due to sufficient evidence of no association *or* due to insufficient evidence of an association. An important feature of those studies, as well as this dissertation, is that all simulated data with at least four indicators per factor. If there are fewer than four indicators per factor, TETRAD returns some information about the model, but might not return the correct structure.

It is important to point out that TETRAD has other functions, including one that simply identifies clusters (FindPattern)- which to contrast with BuildPureClusters could be nicknamed "BuildMessyClusters"- it does not seek anything like simple structure. FindPattern is the TETRAD algorithm that feeds results into BuildPureClusters, and is not implemented specifically in TETRAD (although it can be isolated using the Java source code). A discussion of the FindPattern algorithm is beyond the scope of this work, but an interesting follow up study would be to compare the performances, and interpretabilities, of FindPattern and BuildPureClusters on data that vary along dimensions similar to those studied in this dissertation. Again, the algorithms have completely different purposes and may utilize the same vanishing tetrads as evidence for "decisionmaking" (whether to keep or eliminate associations between variables), but they have totally different outcomes (and FindPattern is the analytic step that feeds results to BPC); and comparing their results would be challenging. The recommendations based on this dissertation work are focused on further exploring the performance of BPC with respect to uncovering dimensionality in data.

5.5.2 Next research steps

These results suggest that BPC could be more useful in ruling zero or one factors *in*to consideration, since AIC from confirmatory analyses was deemed to reflect the correct solution 20-38% of the time for zero factors (i.e., failures to converge were treated as "correct"). Minimum average partials (MAP, Velicer, 1976) or parallel analysis (PA, Horn, 1965) could be more useful than CFA (or using AIC) for this study – an empirical question that would be an interesting follow-up to this dissertation. Comparing MAP and PA to CCA would be a more direct comparison of the dimensional recovery performance of the three methods, and information, like loadings, would *not* need to be sacrificed in order to obtain comparable outcomes from these methods the way it was for CFA results in the present study.

Like Velicer (2000) recommended, the results of this dissertation tend to support a two-step approach to estimating the dimensionality prior to a CFA or confirmatory analysis. Investigators could use BPC to generate the lower bound of the number of factors in the solution, as an alternative to using minimum average partials (MAP, Velicer, 1976) or parallel analysis (PA, Horn, 1965). BPC would be an interesting alternative to MAP and PA because BPC uses the same covariance matrix as will be used in the confirmatory analysis, but computes different estimations (while MAP and PA use the same matrix and the same estimations as the confirmatory analysis, thus potentially creating more sample-specific or overfit results). The second step would be to perform common factor analysis to estimate loadings from the BPC-identified lower bound to the maximum reasonable number of factors (e.g., Bartholomew & Knott, 1999); solution fitnesses can then be compared using AIC or other criteria (Anderson, 2008).

As argued by Velicer (2000) and others, identifying the latent structure that underlies a set of observed variables, and the estimation of factor loadings, requires *both* dimension estimation *and* another algorithm for the loadings. The choice for estimation of dimensions should consider whether zero and one factor solutions are reasonable, and whether causal structure is the ultimate goal. Another follow-up study would be to compare performance of BPC (CCA), MAP and PA at dimension estimation when the data are causally generated and when they are not (i.e., if latent variables emerge from the indicators, rather than causing them).

A condition that was not included in this dissertation is a path model, where the common cause of observed variables-correlation is another observed variable. Since TETRAD (and BPC) are unique in their inclusion of a non-vanishing partial correlations test, to rule out observed variables as causes in the model. Also, an aim of the TETRAD developers is to support uncovering measurement models in support of the development and testing of structural models. For example, the BPC algorithm is the step before the TETRAD algorithm MIMBuild, multi-indicator model builder, which explores the tetrads and partial correlations, as described in Chapter 2, for evidence of structural relations between the measurement models identified by BPC. In this context, the sequential nature of the estimation by CCA might provide different outcomes as compared to another approach to structural equation modeling

(SEM) such as the exploratory SEM methods recently introduced in MPlus (Asparouhov & Muthén, in press).

Finally, the simulations were designed to challenge the capabilities of the two methods (small sample sizes plus weak factors plus correlated factors in some cases); an interesting follow-up study would be to document what parameters each method requires in order to be 100% accurate and consistent. This particular study would be very useful to show, above and beyond the divergent purposes, approaches, and data requirements for these methods, where they function best, respectively, in terms of sample size, strengths of associations between indicators and factors, number of factors, and the other features manipulated in this dissertation.

5.5 Conclusions

This dissertation was the first study of how CCA, as compared to CFA by OLS, performs at identifying the dimensionality of data with simple structural features, varying in terms of specific features of complexity often encountered by in multivariate data: whether factors are correlated, weaker and stronger factor loadings, whether factors share indicators, and when one factor is weaker than the others. It is also the first study to explore accuracy and consistency (invariance), and whether these two characteristics are sample size dependent, for these methods. The research found that CCA and CFA generally perform at similar levels in terms of accuracy (generally high) and consistency (generally low) over repeated trials, and that this performance by both methods was affected by sample size, model complexity, the presence of weak common factors, and correlated vs. independent latent variables. The effects of these conditions tended to be omitting factors and indicators by CCA and a combination of omitting factors and indicators and adding extra factors and indicators in CFA. Thus, choosing between these methods will depend on the type of error the investigator is most interested in avoiding.

Although these methods performed relatively similarly, and responded similarly to the conditions imposed in the simulated data, they have different interpretations and can be used in complementary ways. For example, using CCA to estimate (the lower bound for) the number of latent variables in a system uses the covariance matrix in a different way than CFA (or other estimation procedures); thus, it might be possible to avoid propagating the effects of capitalization on chance if a CCA is followed by a confirmatory method. This lower bound will only be meaningful if there is sufficient information for the investigator in solutions (from CCA) that are unidimensional. This is a question of particular interest that is currently being pursued. However, there are many follow up studies that could be pursued, given that the data requirements of the methods are met (e.g., unidimensional measurement models with four or more indicators for the use of CCA).

Appendix 1: Simulation Table

SIMULATION TABLE. 50 samples in each sample size cell were simulated. One summary (outcome) was computed for each set of 50 trials.

Zero factors			
	N=100	N=300	N=500
One factor			
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	PURE	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	PURE	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	PURE	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	PURE	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 1 indicator	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			

Simulation Table, Cont.

four factors	2f share 1 indicator	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 1 indicator	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 1 indicator	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 2 indicators	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 2 indicators	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 2 indicators	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
four factors	2f share 2 indicators	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			

Simulation Table, con't.

six factors	PURE	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	PURE	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	PURE	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	PURE	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 1 indicator	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 1 indicator	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 1 indicator	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			

Simulation Table, con't.

six factors	2f share 1 indicator	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 2 indicators	no weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 2 indicators	no weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 2 indicators	1 weak factor	independent factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			
six factors	2f share 2 indicators	1 weak factor	correlated factors
	N=100	N=300	N=500
load weak			
load mid			
load strong			

Appendix 2: Complete Scoring Rules

A. Identifying best solution - the main outcome of the study

The CCA (TETRAD) results always returned only one model. CEFA analyses were done for the four and six factor data, extracting from 1 – true+2 number of factors. The best solution had to be identified.

A1. CEFA analyses were run extracting 1-6 (from 4 factor data) or 1-8 (from 6 factor data) factors. Which of these 6 or 8 solutions is the best one? "best" is determined by AIC.

A2. the over 68,000 cefa results will be reduced to 50 from each of the conditions involving 4 and 6 factors, which is 10,800 of the 11,400 total simulations.

A3. the number of factors identified by AIC as the best solution is a data point - for each of the 50 trials in each condition, retain number of factors identified.

B. characterizing the best solution - extracting the other raw data needed for this study.

B1. In the best solution from CEFA, what indicators loaded on what factors? A data file was created for the tetrad results that had variables called "latent1" - 1=latent variable 1 was found, 0=latent variable 1 was not found. This isn't the "first factor", it is just whether or not the solution found ONE factor. For four factor data, the files had Latent1, latent2, ..., latent6, and 6 latent variables were extracted by tetrad about four times over all the trials. Similarly, for six factor data, 8 factors were extracted by tetrad one time; 7 factors were extracted about 40 times overall.

B1a. Each extracted factor was coded whether it was the factor made of items:

- 1-4, 5-8, 9-12, or 13-16 for f4.s0 (four factors in true population, no items shared by latents)
- 1-4, 5-8, 9-11, or 12-15 for f4.s1 (four factors in true population, one item (x8) shared by 2 latents)
- 1-4, 5-8, 9-10, or 11-14 for f4.s2 (four factors in true population, two items (x7,x8) shared by 2 latents)
- 1-4, 5-8, 9-12, 13-16, 17-20, 21-24 for f6.s0 (six factors in true population, no items shared by latents)
- 1-4, 5-8, 9-11, 12-15, 16-19, 20-23 for f6.s1 (six factors in true population, one item (x8) shared by 2 latents)
- 1-4, 5-8, 9-10, 11-14, 15-18, 19-22 for f6.s2 (six factors in true population, two items (x7, x8) shared by 2latents)

B1b. Each factor had a 'sentinel' - an indicator which, if observed, tells which latent variable has been found. The sentinels were always the same:

- 1, 5, 9, 13 for f4.s0 (the f4.s0 ending reflects four factors in true population, no items shared by latents)
- 1, 5, 9,12 for f4.s1 (four factors in true population, one item (x8) shared by2 latents)
-1, 5, 9, 11 for f4.s2 (four factors in true population, two items (x7, x8) shared by 2 latents)

- 1, 5, 9, 13, 17, 21 for f6.s0 (six factors in true population, no items shared by latents)

- 1, 5, 9, 12, 16, 20 for f6.s1 (six factors in true population, one item (x8) shared by2 latents)

- 1, 5, 9, 11, 15, 19 for f6.s2 (six factors in true population, two items (x7, x8) shared by 2 latents)

B1c. Indicators are said to have loaded on any of the found factors in the winning CEFA solution if that indicator's loading/SE(loading)>1.96 [loadings after geomin rotation]

B1.d Indicators can load on >1 factor in CEFA solutions.

B2. The tetrad results from this step represent, for each of 50 solutions, the file name generating the results, and then the indicators, entered as the observed data for the variables like "loader1.1", which is that indicator which loaded on factor 1 as indicator #1, and "loader 2.3", which is that indicator which loaded on factor 2 as indicator #3. The data were entered for tetrad results as string variables, "x1" or "x12". Sorting was subsequently found to be 10 times faster if the xs were removed and the variables treated as numbers instead.

B2a. If a factor that should have been found was not found, e.g., factor 3 in either 4 or 6 factor data or factor 5 in the 6 factor data and not the 4 factor data, then the observation for that variable (e.g., loader3.2, loader 5.1) was missing.

C. Summarizing the raw data: counting errors, secondary outcomes.

The raw tetrad output (described in B1 and B2 above) was entered manually from the tetrad analyses. These files comprised the 450 trials for one set of conditions, e.g., f4.s1.w1.i0 (four factors in population (f4), two of the 4 latents share one indicator (s1), one of the factors has loadings 50% of the others (a weak factor present, w1) and the four factors correlate at .5 (factors not independent, i0). This is 50 trials at each of three loading strengths (weak LW, medium LM, strong LS) for 3 sample sizes (N100, N300, N500) per strength.

C1. The primary outcome of correct # of latent variables found was obtained for each of the 50 trials as the sum of the variables "latent1", "latent2", etc, up to latent8 (i.e., if too many factors were found, they were still counted). Then four summarizing variables were computed.

C1a. The number of times latent4 or latent6 was found, divided by 50, the number of trials gives the number of times <at least the correct number was achieved>/<opportunities to find factors>: *dimdisc*.

C1b. the number of factors in each solution was computed: *found* (how many factors were found in each solution)

C1c. omissions of factors were identified as 1-, 4-, or 6-found: *fac_om*.

C1d. the number of extra factors found was identified as 4- or 6-found>0.

The mean and SD of these four variables over 50 trials was recorded in a new file, in addition to other variables computed after the results from item C2 below were completed.

C2. For every one of the 50 trials for the specific combination of conditions, *each factor* was assessed separately for errors relating to indicators specifically. To do this, the data were sorted using the information from a single factor, e.g., loader1.1, loader1.2, ..., loader1.7. Once the data were sorted based on the sentinel for whatever was recorded as having loaded on the first factor to be extracted (could have been the items of "factor 1", x1-x4 or those for "factor 4", x12-x16), the set of 50 factor 1 results were printed. Errors were identified, classified, and counted as follows: In the tetrad results, the sentinel was the first loader in about 70% of cases. x1 and x13 (and x16, x21) were most often the first loader; x5 was usually, but not always first, and x9 was usually NOT first, but was almost always last whenever it appeared.

note: omission or commission errors involving a sentinel were not treated any differently (counted more or less) then those for non-sentinels - sentinels were only used for decisionmaking in terms of counting errors.

C2a. Omission errors: any time a sentinel was observed without one or more of the 3 other nonsentinels it was supposed to be with, each 'omitted' non-sentinel is a single omission error for that factor. When a majority of loaders on one factor, or a single sentinel, was not observed, arbitrary–but consistently applied– decisions were made that led to the same number and characterizations of the errors whether one or another factor was 'identified.

C2a. 1) if a sentinel is observed with 3 other indicators that are not part of the sentinel's group, those are 3 omissions as well as 3 commissions - each error counts twice.

C2a. 2) if two sentinels plus two correct indicators are observed on one factor, this is counted as two omissions and two commissions. The first sentinel was chosen to represent to which factor the two omission and two commission errors were 'charged'.

C2a. 3) if one sentinel, two or three correct indicators and a second sentinel were observed, it was one omission and one commission error (the factor being identified as that constituted by the sentinel and its indicators).

C2a.4) if there were equal numbers of nonsentinels, e.g., x2, x3, x6, x7, the first factor suggested (f1 in this case) was ruled the factor and the other two indicators were called commissions.

C2a.5) if there were more nonsentinels from one factor than another, majority ruled to determine which were commissions/omissions.

C2a.6) errors could be just omissions, e.g., if one sentinel and one of its nonsentinels were put on a factor together, with nothing else, it was two omissions.

C2a.7) errors could be just commissions, e.g., if >4 indicators were assigned to a factor, this could represent a complete factor (1 sentinel +3 non sentinels correctly assigned) plus one indicator misassigned.

C2a. 8) in several solutions a factor had only sentinel indicators (i.e., no non-sentinels loading on it), e.g., x1, x5, x9, x12. This was assigned to x1 (factor 1) with 3 omission and 3 commission errors.

C2b. Commission errors. Any time an indicator was placed on a factor where its sentinel did not appear (e.g., x1, x2, ,x3, x4, x6), OR, if it was placed with its sentinel but as the minority (e.g., x1, x2, x3, x5, x6), OR with other of its nonsentinel co-loaders but as the minority on a factor (e.g., x2, x3, x4, x6, x7), it was counted as a commission error.

C2c. For shared indicator conditions:

C2c. 1) whenever x7 (S2 only), or x8 (S1 or S2) or both (S2 only) were observed to load together without either x5 or x9, the two sentinels for the factors by which these indicators were shared, if they were the majority (or only indicators), they were assigned to the factor with sentinel x5.

C2c.2) for results like x5, x7, x8, x9 - which is ambiguous as to whether it is factor 2 (x5 sentinel) or factor 3 (x9 sentinel), it was ruled factor 2 and the x9 loading was called an omission (x6 is missing from factor 2, which was x5, x6, x7, x8) AND a commission (x9 is not part of factor 2).

C2c. 3) similarly, for outcomes x10, x6, x7, x8 - there is no sentinel and in S2 conditions, x7 and x8 are shared by the factor with x6 AND the factor with x10 - so this is an ambiguous result. It was assigned to factor 2 (6,7,8) with x10 as the omission and commission error.

D. Each factor of the winning solution was evaluated for errors separately. the full solution (e.g., all four factors and all their indicators) were not examined simultaneously as "a solution" because how to do this consistently using the tetrad output was unknown.

D.1. For each factor, the number of times, of the 50 trials, that the factor was found, the per factor omission rate was computed as follows:

D.1a. for the factor extracted first, omission errors were added from those times where the first-extracted factor was made of items 1-4, those where the first-extracted factor was made of items 5-8, those where the first-extracted factor was made of items 9-12, those where the first-extracted factor was made of items 13-16, and so on for the 6 factor solution.

D.1.a. 1) in addition to these *recorded* omission errors, there were unrecorded omission errors: every time the factor was NOT found, it counted as 4 omission errors. 50- number of times a factor was found = number of times factor was not found.

D.1b. the formula for the number of omission errors for the first-extracted factor (F1) was:

{[4 x (50-<number of times F1 was found>)] + (omissions when F1 was x1-x4+omissions when F1 was x5-x8+omissions when F1 was x9-x12+omissionswhen F1 was x13-x16)}/200

There are 200 opportunities for omission errors on a given factor, because every factor had 4 indicators and every factor was searched for in 50 trials (4x50=200).

D.2 Commission errors were a similar sum of commissions counted when the particular factor was made up of items from each possible factor, like for omissions. However, the denominator for commission errors was the total number of times the factor was found. That is, D.2a. the formula for the number of commission errors for the first-extracted factor (F1) was:

(commissions when F1 was x1-x4+commissions when F1 was x5-x8+commissions when F1 was x9-x12+commissions when F1 was x13-x16)}/total number of times F1 was found.

D.3. Each condition, then, yielded 5-8 omission 'scores' (D1), 5-8 commission scores (D2), plus the four results describing the number of latent variables found (C1a - C1d), two of which had SDs (avg number of factors omitted and average number of factors COMmitted). With these summaries (summarizing the 50 trials for each condition set), the features the conditions represent (number of latents in the population, presence of weak factor, independence, purity, loadings and N) were combined into a single file.

Bibliography

Ackerman TA, Gierl MJ & Walker CM. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. Educational Measurement: Issues and Practice 22: 37-53.

Afifi A, Clark VA & May S. (2004). Computer-aided multivariate analysis 4E. Boca Raton, FL: Chapman & Hall.

Akaike H. (1973). Information theory and an extension of the maximum likelihood principle. In (BN Petrov, F Csaki, Eds), 2nd International Symposium on Information Theory. Akademiai Kiado, Budapest. Pp. 267-281.

Anderson DR. (2008). Model Based Inference in the Life Sciences: A primer on evidence. New York, NY: Springer.

Anderson DR, Burnham KP, White GC. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. Journal of Applied Statistics 25(2): 263-82.

Asparouhov T & Muthén B. (in press). Exploratory structural equation modeling. Structural Equation Modeling. Downloaded from http://www.statmodel.com/papers.shtml 10 March 2009.

Bandolos D. (1996). Confirmatory factor analysis. In Stevens J. Applied Multivariate Statistics for the Social Sciences, 3E. Mahwah, NJ: Lawrence Erlbaum Associates. Pp. 389-420.

Bandalos DL & Boehm MR. (in press). Four common misconceptions in exploratory factor analysis. In CE Lance & RJ Vandenberg (Eds.), Statistical and Methodological Myths and Urban Legends: Received Doctrine, Verity, and Fable in the Organizational and Social Sciences. Mahwah, NJ: Erlbaum.

Bartholomew DJ. (2007). Three faces of factor analysis. In R. Cudeck & RC MacCallum (Eds.) Factor Analysis at 100: Historical developments and future directions. Mahwah, NJ: Lawrence Erlbaum Associates. Pp. 9-21.

Bartholomew DJ, Knott M. (1999). Latent Variable Models and Factor Analysis. London: Arnold.

Bentler PM. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In S-Y Lee (Ed.), Handbook of Latent Variable and Related Models. Amsterdam: Elsevier. Pp. 1-19.

Bollen KA. (1990). Outlier screening and a distribution-free test for vanishing tetrads. Sociological Methods and Research 19:80-92.

Bollen KA. (1989). Structural Equations with Latent Variables. New York, NY: Wiley.

Bollen KA. (1996). A limited information estimator for LISREL models with or without heteroscedastic errors. In GA Marcoulids and RE Schumacker (Eds.), Advanced Structural Equation Modeling. Mahwah, NJ: Erlbaum. Pp. 227-241.

Bollen KA, Ting K. (1993). Confirmatory tetrad analysis. Sociological Methodology 23: 147-175.

Bollen and Ting (2000). A tetrad test for causal indicators. Psychological Methods 5(1): 3-22.

Borsboom D. (2005). Measuring the mind. Cambridge, UK: Cambridge University Press.

Bozdogan H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 52, 345-70.

Briggs NE & MacCallum RC. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. Multivariate Behavioral Research 38(1): 25-56.

Brown SC & Craik FIM. (2000). Encoding and retrieval of information. In E. Tulving and FIM Craik (Eds.) The Oxford Handbook of Memory. New York NY: Oxford University Press. Pp. 93-106.

Browne MW. (2001). An overview of analytic rotation in exploratory factor analysis. Multivariate Behavioral Research 36(10): 111-150.

Browne MW & Cudeck R. (1993). Alternative ways of assessing model fit. In, KA Bollen, SJ Long (Eds), Testing structural equations. Newbury Park, CA: Sage.

Browne, M. W., Cudeck, R., Tateneni, K. & Mels G. (2008). CEFA: Comprehensive Exploratory Factor Analysis, Version 3.02 [Computer software and manual]. Retrieved 18 February 2009 from http://faculty.psy.ohio-state.edu/browne/

Burnham KP & Anderson DR. (2002). Model selection and multimodel inference, 2E. New York, NY: Springer.

Carroll JB. Human Cognitive Abilities: A survey of factor-analytic studies. (1993). Cambridge, UK: Cambridge University Press.

Cattell RB. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press.

Cowles M. (2001). Statistics in Psychology: An Historical Perspective, 2E. Hillsdale, NJ: Lawrence Erlbaum Associates.

Corter JE. (1995). Using clustering methods to explore the structure of diagnostic tests. In PD Nichols, SF Chipman & RL Brennan (Eds), Cognitively Diagnostic Assessment. Hillsdale, NJ: Lawrence Erlbaum Associates. Pp. 305-326.

Cudeck, R. (2000). Exploratory factor analysis. In HEA Tinsley & SD Brown (Eds.), Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego, CA: Academic Press. Pp. 265-296.

Cureton EE & D'Agostino RB. (1983). Factor Analysis: An Applied Approach. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dilalla L. (2000). Structural equation modeling: Uses and issues. In Tinsley and Brown (Eds) Handbook of Multivariate Statistics and Multivariate Modeling. San Diego, CA: Academic Press. Pp. 439-464.

Drton M, Sturmfels B & Sullivant S. (2007). Algebraic factor analysis: tetrads, pentads and beyond. Downloaded from http://front.math.ucdavis.edu/math.ST/0509390 25 September 2007.

French, BF, Finch WH. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. Structural Equation Modeling 13(3): 378-402.

Gorsuch RL. (1983). Factor Analysis, 2E. Hillsdale, NJ: Lawrence Erlbaum Associates.

Glymour C, Scheines R, Spirtes P & Kelly K. (1987). Discovering Causal Structure. Academic Press: San Diego, CA.

Harman HH. (1976). Modern Factor Analysis, 3E (revised). The University of Chicago Press: Chicago, IL.

Hayduk LA & Glaser DN. (2000). Jiving the four-step, waltzing around factor analysis, and other serious fun. Structural Equation Modeling 7(1):1-35.

Herting JR & Costner HL. (2000). Another perspective on "the proper number of factors" and the appropriate number of steps. Structural Equation Modeling 7(1): 92-110.

Horn JL. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika 30: 283-300.

Hoyle R. (2000). Confirmatory factor analysis. In Tinsley and Brown (Eds) Handbook of Multivariate Statistics and Multivariate Modeling San Diego, CA: Academic Press. Pp. 465-497

Hoyle RH & Duvall JL. (2004). Determining the number of factors in exploratory and confirmatory factor analysis. In: D. Kaplan (Ed.) The Sage handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage. Pp. 301-315.

Jennrich RI. (2007). Rotation altorithms: From beginning to end. In S-Y Lee (Ed.) Handbook of Computing and Statistics with Applications, Vol 1: Handbook of Latent Variable and Related Models. Amsterdam, The Netherlands: North-Holland. Pp..45-63.

Johnson RA & Wichern DW. (2002). Applied multivariate statistical analysis, 5E. Upper Saddle River, NJ: Prentice Hall.

Jöreskog KG. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 34(2):183-202.

Jöreskog KG. (2007). Factor analysis and its extensions. In R. Cudeck & RC MacCallum (Eds.) Factor Analysis at 100: Historical developments and future directions. Mahwah, NJ: Lawrence Erlbaum Associates. Pp. 47-77.

Kano Y. (2007). Selection of manifest variables. In S-Y Lee (Ed.), Handbook of Latent Variable and Related Models. Amsterdam: Elsevier. Pp 65-86.

Kelly T. (1928). Crossroads in the Mind of Man. Stanford, CA: Stanford University Press.

Kenny DA. (1979). Correlation and Causality. New York, NY: Wiley.

Kline P. (1994). An easy guide to factor analysis. London, UK: Routledge.

Kline RB. (2005). Principles and Practice of Structural Equation Modeling, 2E. The Guilford Press.

Konishi S, Kitagawa G. (2008). Information Criteria and Statistical Modeling. New York, NY: Springer.

Lawley DN & Maxwell AE. (1962). Factor analysis as a statistical method. London, UK: Butterworths.

MacCallum RC, Widaman KF, Preacher KJ, Hong S. (2001). Sample size in factor analysis: the role of model error. Multivariate Behavioral Research 36(4): 611-637.

Manly BFJ. (2005). Multivariate statistical methods: A primer 3E. Boca Raton, FL: Chapman & Hall.

Maruyama, GM. (1998). Basics of structural equation modeling. Thousand Oaks, CA: Sage.

McQuarrie ADR & Tsai C-L. (1988). Regression and Time Series Model Selection. World Scientific, London, UK.

Millsap RE and Meredith W. (2007). Factorial invariance; historical perspectives and new problems. In, R Cudeck, RC MacCallum (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Earlbaum: Mahwah, NJ. Pp. 131-175.

Mulaik SA. (1972). The foundations of factor analysis. New York: McGraw-Hill.

Mulaik SA. (1986). Factor analysis and Psychometrika: Major developments. Psychometrika: 5191): 23-33.

Nesselroade J. (1994). Exploratory factor analysis with latent variables and the study of processes of development and change. In von Eye and Clogg (Eds), Latent Variables Analysis: Applications for Developmental Research. Thousand Oaks, CA: Sage. Pp.131-154.

O'Connor BP. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods, Instruments and Computers 32(3): 396-402.

Osborne JW, Costello AB & Kellow JT. (2008). Best practices in exploratory factor analysis. In JW Osborne (Ed.) Best Practices in Quantitative Methods. Thousand Oaks, CA: Sage. Pp. 86-99.

Pearl J. (2000). Causality. Cambridge, UK: Cambridge University Press.

Pett MA, Lackey NR, Sullivan JJ. (2003). Making Sense of Factor Analysis: The use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage.

Preacher KJ, MacCallum RC. (2002). Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. Behavior Genetics 32(2): 153-161.

Preacher KJ & MacCallum RC. (2003). Repairing Tom Swift's electric factor analysis machine. Understanding Statistics 2(1): 13-43.

Reyment RA & Jöreskog KG. (1993). Applied Factor Analysis in the Natural Sciences. Cambridge, UK: Cambridge University Press.

Royce JR. (1963). Factors as theoretical constructs. In DN Jackson & S Messick (Eds.) Problems in Human Assessment. New York: McGraw-Hill.

Scheines R, Glymour C & Spirtes P. (2005). The TETRAD project: Causal Models and Statistical Data (freeware & Manual/Help files). Downloaded 1Jan 2006 from http://www.phil.cmu.edu/projects/tetrad/

Scheines R, Spirtes P, Glymour C, Meek C & Richardson, T. (1998). The TETRAD Project: constraint-based aids to causal model specification. Multivariate Behavioral Research 33(1): 65-117.

Schwarz G. (1978). Estimating the dimension of a model. Annals of Statistics 6, 461-4.

Sheskin DJ. (2004). Handbook of Parametric and Nonparametric Statistical Procedures, 3E. Boca Raton, FL: Chapman & Hall.

Silva, R. (2005). Automatic Discovery of Latent Variable Models. Unpublished PhD Thesis. Machine Learning Department, Carnegie Mellon University. Downloaded at http://www.statslab.cam.ac.uk/~silva/ June 2007.

Silva, R.; Scheines, R.; Glymour, C. and Spirtes, P. (2006). Learning the structure of linear latent variable models. Journal of Machine Learning Research 7(Feb):191—246.

Skrondal A & Rabe-Hesketh S. (2004). Generalized Latent Variable Modeling: Multilevel, longitudinal and structural equation models. Boca Raton: Chapman & Hall/CRC.

Spearman C. (1904). General intelligence objectively determined and measured. American Journal of Psychology 15: 201-293.

Spearman C. (1927). The Abilities of Man. New York, NY: Macmillan.

Spirtes P, Glymour C & Scheines R. (2000). Causation, Prediction and Search, 2E. The MIT Press: Cambridge, MA.

Stevens JP. (2002). Applied multivariate statistics for the social sciences 4E. Mahwah, NJ: Lawrence Erlbaum Associates.

Thurstone LL. (1937). Vectors of the Mind. Chicago, IL: University of Chicago Press.

Thurstone LL. (1940). Experimental study of simple structure. Psychometrika 5: 153-168.

Thurstone LL. (1947). Multiple factor analysis. Chicago, IL: University of Chicago Press.

Velicer WF. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika 31: 321-327.

Velicer WF, Eaton CA & Fava JL. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In RD Goffin, E Helmes (Eds.) Problems and Solutions in Human Asssessment: Honoring Douglas N. Jackson at Seventy. Norwell, MA: Kluwer Academic Publishers, pp. 41-71.

Verma T & Pearl J. (1988). Causal networks: Semantics and expressiveness. In *Proceedings of the 6th Conference on Uncertainty in Artificial Intelligence*, vol 4. pp 69-76. Amsterdam: Elsevier.

Wager C, Vaida F, Kauermann G. (2007). Model selection for penalized spline smoothing using Akaike information criteria. Australian & New Zealand Journal of Statistics 49(2) 173-90.

Wang W-C, Wilson M, Adams RJ. (1997). Rasch models for multidimensionality between items and within items. In M Wilson, G Engelhard and K Draney (Eds.). Objective Measurement: Theory into Practice, Vol 4. Greenwich, CT: Ablex. Pp. 139–155

Wickens TD. (1994). The Geometry of Multivariate Statistics. Mahwah, NJ: Lawrence Erlbaum Associates.

Widaman KF. (2007). Common factors versus components: Principles and principals, errors and misconceptions. In, R Cudeck, RC MacCallum (Eds.), Factor Analysis at 100: Historical Developments and Future Directions. Mahwah, NJ: Lawrence Earlbaum. Pp. 177-203.

Yates A. (1987). Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis. Albany, NY: State University of New York Press.

Zhang J, Stout W. (1999). The theoretical detect index of dimensionality and its application to approximate simple structure. Psychometrika 64(2): 213-249.