CAUTIONS ON USING MODEL FIT TO CHOOSE NUMBER OF FACTORS IN EFA

AMANDA KAY MONTOYA
DEPARTMENT OF PSYCHOLOGY
THE OHIO STATE UNIVERSITY

MICHAEL EDWARDS
DEPARTMENT OF PSYCHOLOGY
ARIZONA STATE UNIVERSITY

Slides available at github.com/akmontoya/NCME2017

Overview

- Factor Analysis and Scale Development
- Current Practices in Selecting Number of Factors in EFA
- Using Model Fit to Select Number of Factors
- Simulation: Can we use model fit to select number of factors?
  - Methods
  - Results
- Discussion

"[My answer] is not an attempt to say how many factors there are in the universe of psychological content – for the answer to this theoretical question is obviously that the number of factors in any psychological domain is infinite. I am just presenting an answer which will provide one with a basis for finding that number of factors which may be yielded **necessarily**, **reliably,** and **meaningfully** – from the data at hand."

Henry F. Kaiser

"The application of electronic computers to factor analysis" 1960

Factor Analysis and Scale Development

Scale development is an important part of education research and research in other academic fields.

Factor analysis is a very common analytical method in scale development.

- Dimensionality
- Reliability
- Validity

Dimensionality: Exploratory Factor Analysis (EFA) is used to determine the *number of factors* which can adequately represent the data.

A variety of methods have been proposed for determining the number of factors using EFA

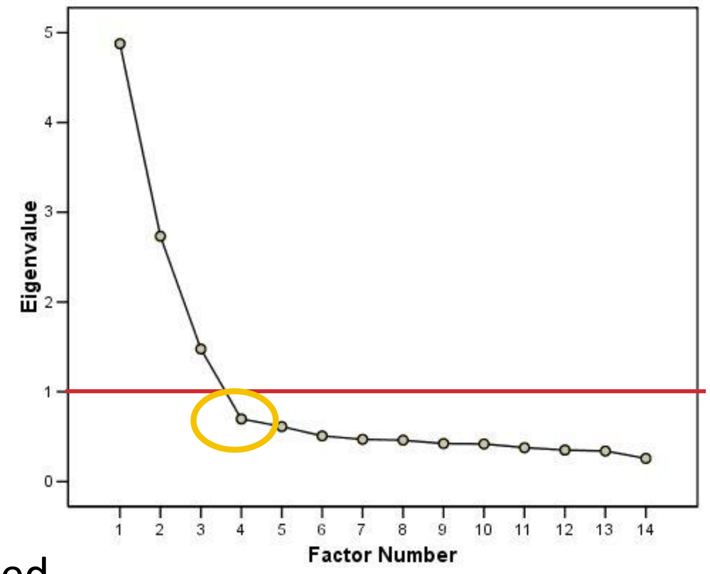
- Recent model fit indices (e.g. RMSEA) have been proposed to address this problem
- We believe this practice could be problematic

Methods for Selecting Number of Factors

Number of Eigenvalues greater than 1 (Kaiser, 1960)

Scree Plot (Cattell, 1966)

Select number of eigenvalues



Minimum Average Partial (MAP) (Velicer, 1976)

Stepwise partialing of components using PCA.
Select number of factors who’s extraction resulted in the lowest residual average squared partial correlation.

Parallel Analysis (Horn, 1965)

Simulation based approach to estimating the distribution of eigenvalues under zero factor model, then select number of factors based on the simulated distribution

Goodness of Fit Statistics

Recommended by Asparouhov & Muthén, 2009; Barendse, Oort & Timmerman, 2015; Floyd & Widaman, 1995; Preacher, Zhang, Kim, & Mels, 2013

Using Model Fit to Select Number of Factors

General Method:

Successively fit models with additional factors until model fit index falls below or above some specific cut off

Fit Indices and potential cut-offs (Hu and Bentler, 1998):

- Root Mean Square Error of Approximation (RMSEA) (< 0.05 or 0.08)
- Standardized Root-Mean-Square Residual (SRMR) (< 0.08)
- Tucker-Lewis Index (TLI)/ Comparative Fit Index (CFI) (> 0.9 or 0.95)

This method is becoming more popular as software has begun providing fit indices for exploratory factor analysis models, and the procedure feels similar to a confirmatory model.

Garrido, Abad, and Consoda (2016) showed some fit indices worked well but no better than current methods in simulation.

Using Model Fit to Select Number of Factors

Some issues:

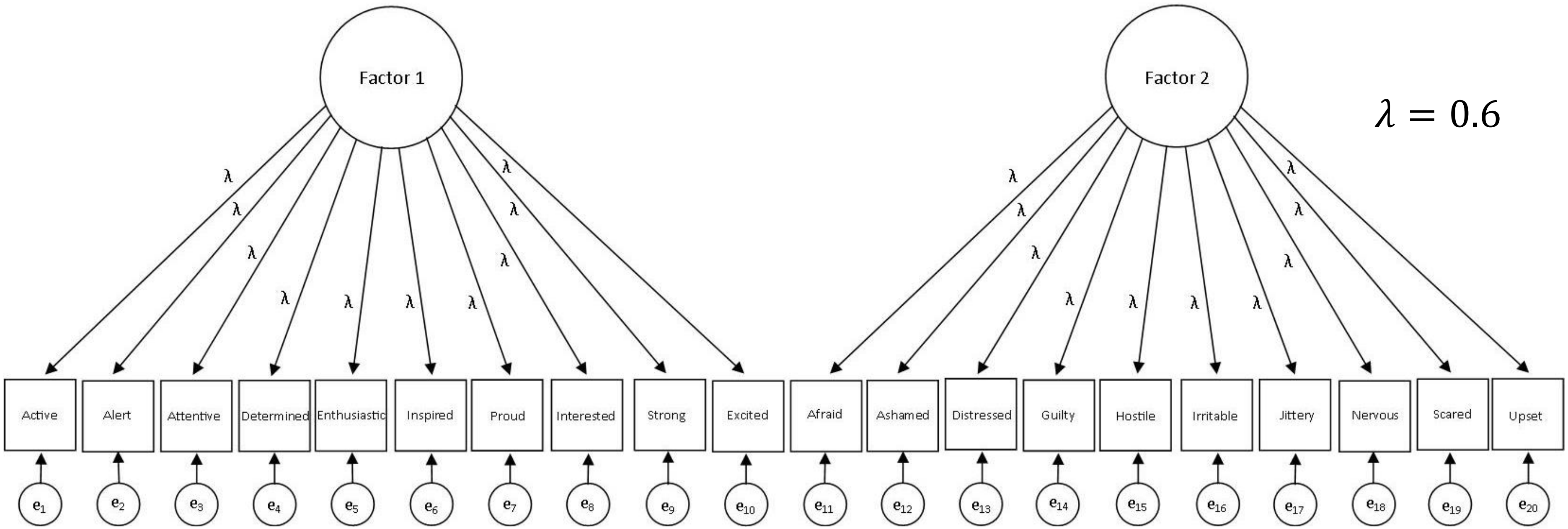
Research in SEM has already called into question the use of cutoffs for model fit in decision making (Chen, Curran, Bollen, Kirby, & Paxton, 2008; MacCallum, Widaman, Zhang, & Hong, 1999; Marsh, Hau, & Wen, 1999; Nye & Drasgow, 2010)

Many other properties of a scale can impact fit indices:

- General fit of the model at the population level (MacCallum, Widaman, Preacher, & Hong, 2001)
  - Garrido, Abad, and Consoda (2016) simulated under perfect model fit
- Correlated residuals
  - A fact of life in scale development
  - We typically don’t consider these in the exploratory phase of scale development

Simulation: Generating Model

Positive and Negative Affect Schedule



$\lambda = 0.6$

Based on model from Ortuño-Sierra, Santarén-Rosell, de Albéniz, & Fonseca-Pedrero, 2015

Simulation: Adding Non-Specific Error

Cudeck and Browne (1992) proposed a method for generating covariance matrices which add *non-specific error* to the matrix.

- Population model does not fit exactly
- Can control the degree of misfit (specify discrepancy δ)
- Maximum likelihood parameter estimates are unchanged

Two-factor structure will not hold exactly in the population. We wanted to replicate the *real life experience* of fitting a factor model, but with simulated data.

Selected degree of misfit based on per-degree of freedom measures of discrepancy

$$\epsilon_p = \sqrt{\frac{\delta_p}{df}}$$

Simulation: Experimental Conditions

100 simulated samples from each of 252 conditions

7 (Type of Correlated Residuals) X 6 (Magnitude of Correlated Residuals) X 6 (Degree of Nonspecific Error)

Type of Correlated Residuals (CR)

- 1 cross-factor CR
- 1, 2, 3, and 4 within-factor CR
- 5 CRs (4 within and 1 cross)

Magnitude of Correlated Residuals

- 0, 0.1, 0.2, 0.3, 0.4, 0.5
- All correlated residuals are the same magnitude

Degree of Nonspecific Error

- ϵ = 0, 0.02, 0.04, 0.06, 0.08, 0.10

Simulation: Data Generation and Analysis

Generation:

Generated 100 samples with N = 1000 on 20 standard normal variables

Used perturbed correlation matrix and generated normal deviates to generate sample correlation matrices for each condition

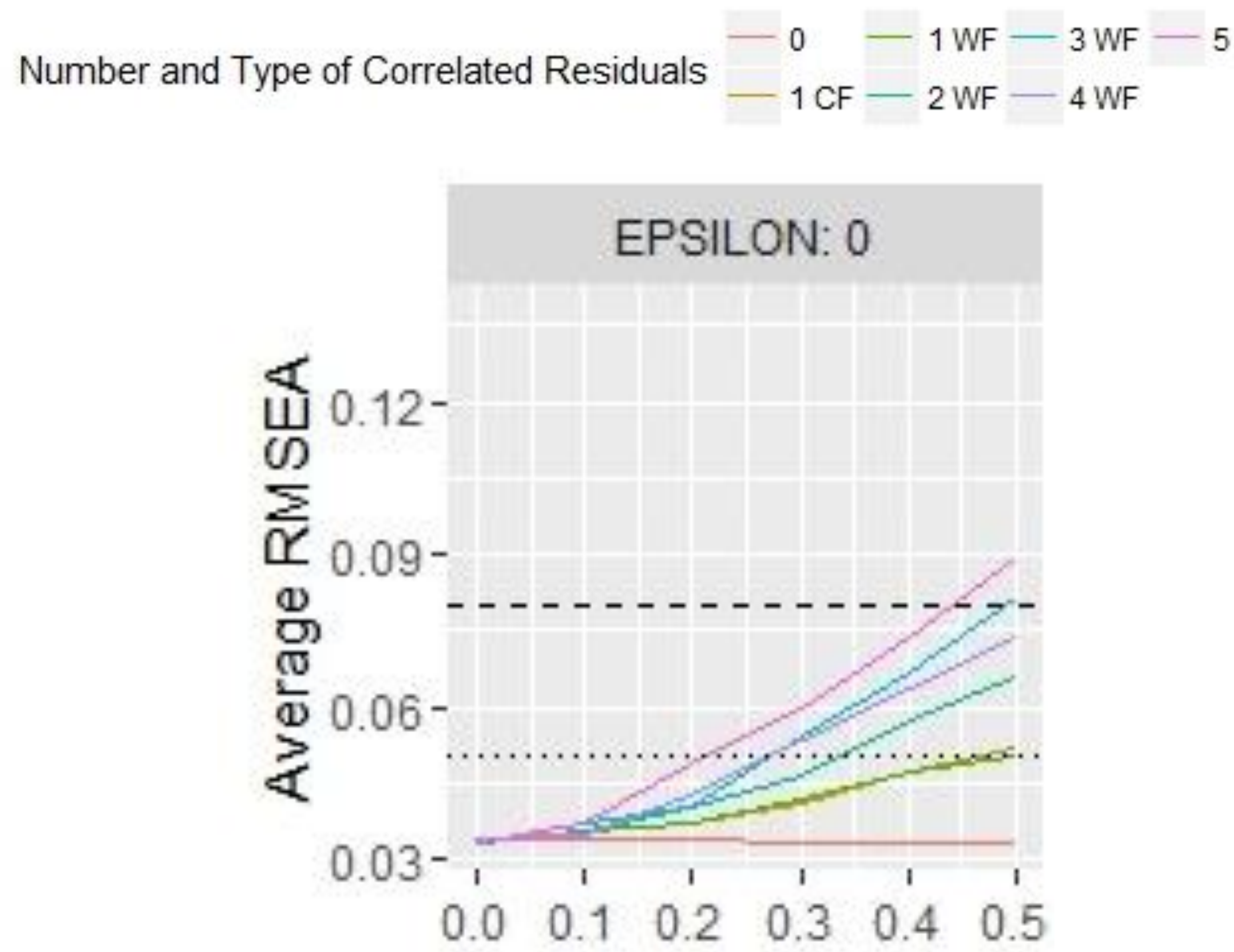
Analysis:

Used Mplus to fit a two-factor EFA (did not include correlated residuals)

Saved fit statistics as outcomes

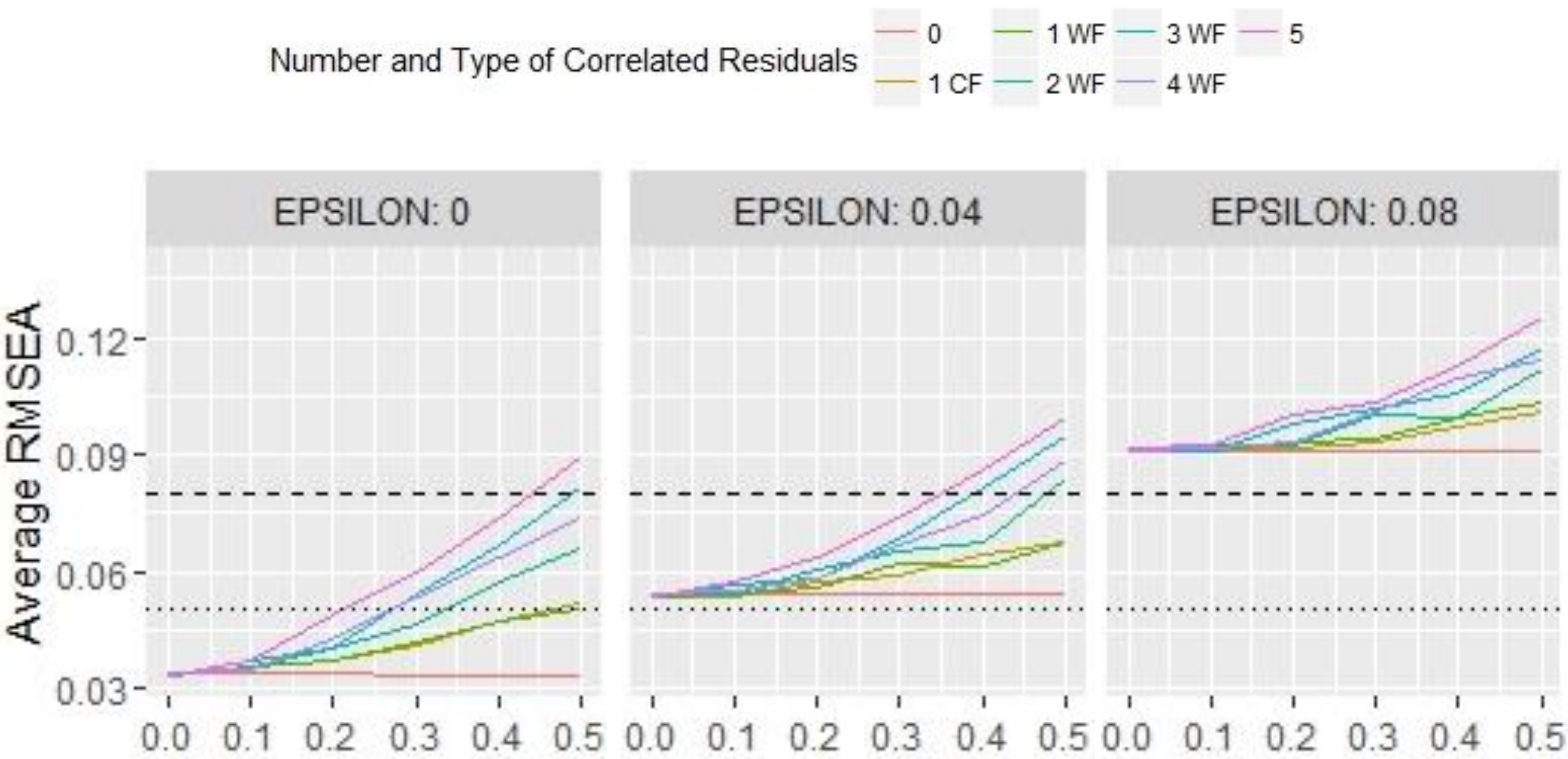
- RMSEA: measure of discrepancy per degree of freedom $\epsilon_s = \sqrt{\frac{\delta_s}{df}}$
- SRMR: Standardized difference between observed and predicted correlation
- TLI: χ^2 based statistic. Fit of 2-factor model to null model, relative to perfect model compared to null model, normed based on degrees of freedom.
- CFI: Similar to TLI with a different penalty for degrees of freedom (penalty of 1 for each df)

Simulation: Results



Magnitude of Correlated Residuals

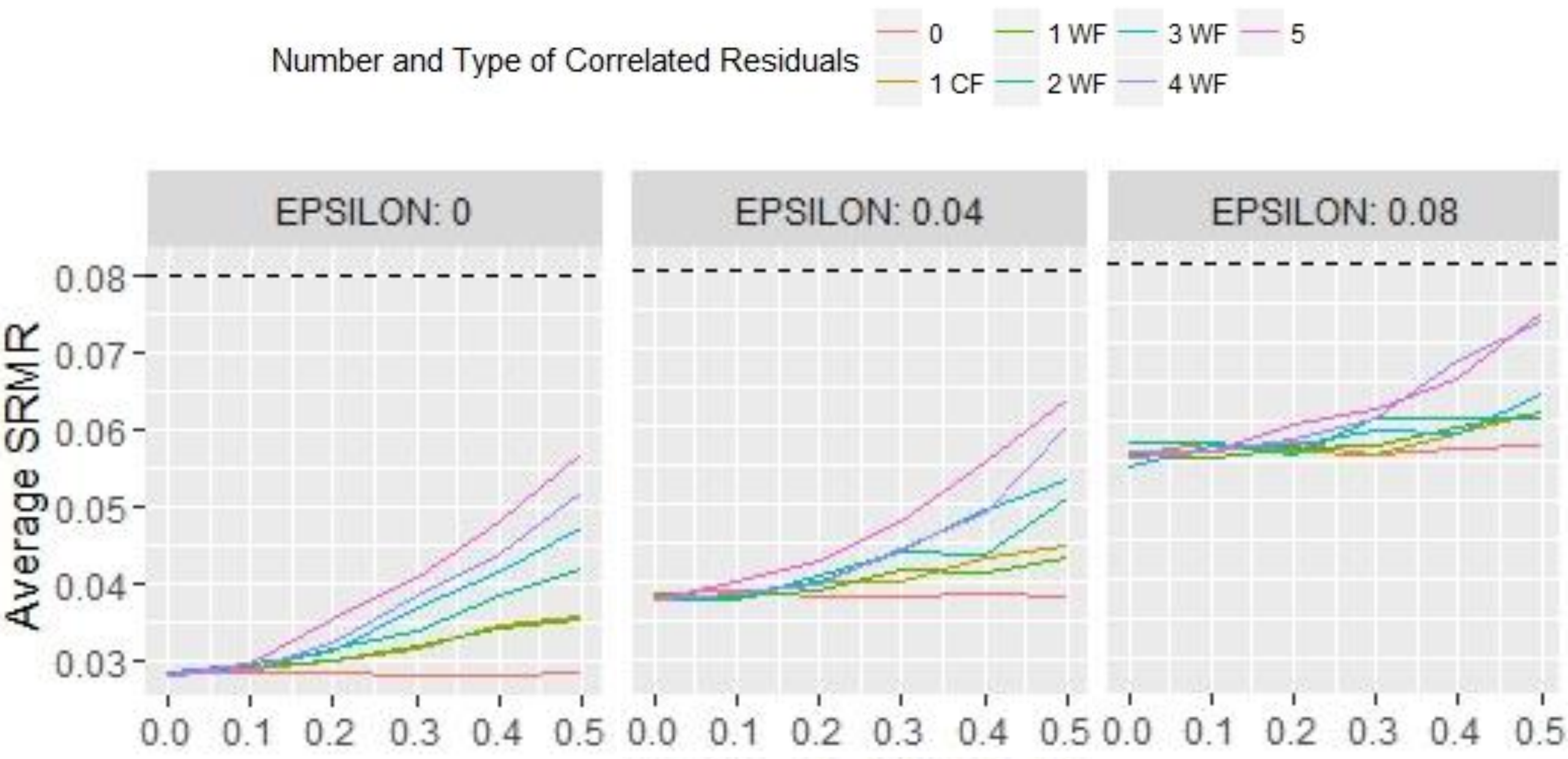
Simulation: Results RMSEA



Magnitude of Correlated Residuals

ϵ = 0.2, 0.6, 0.10 omitted

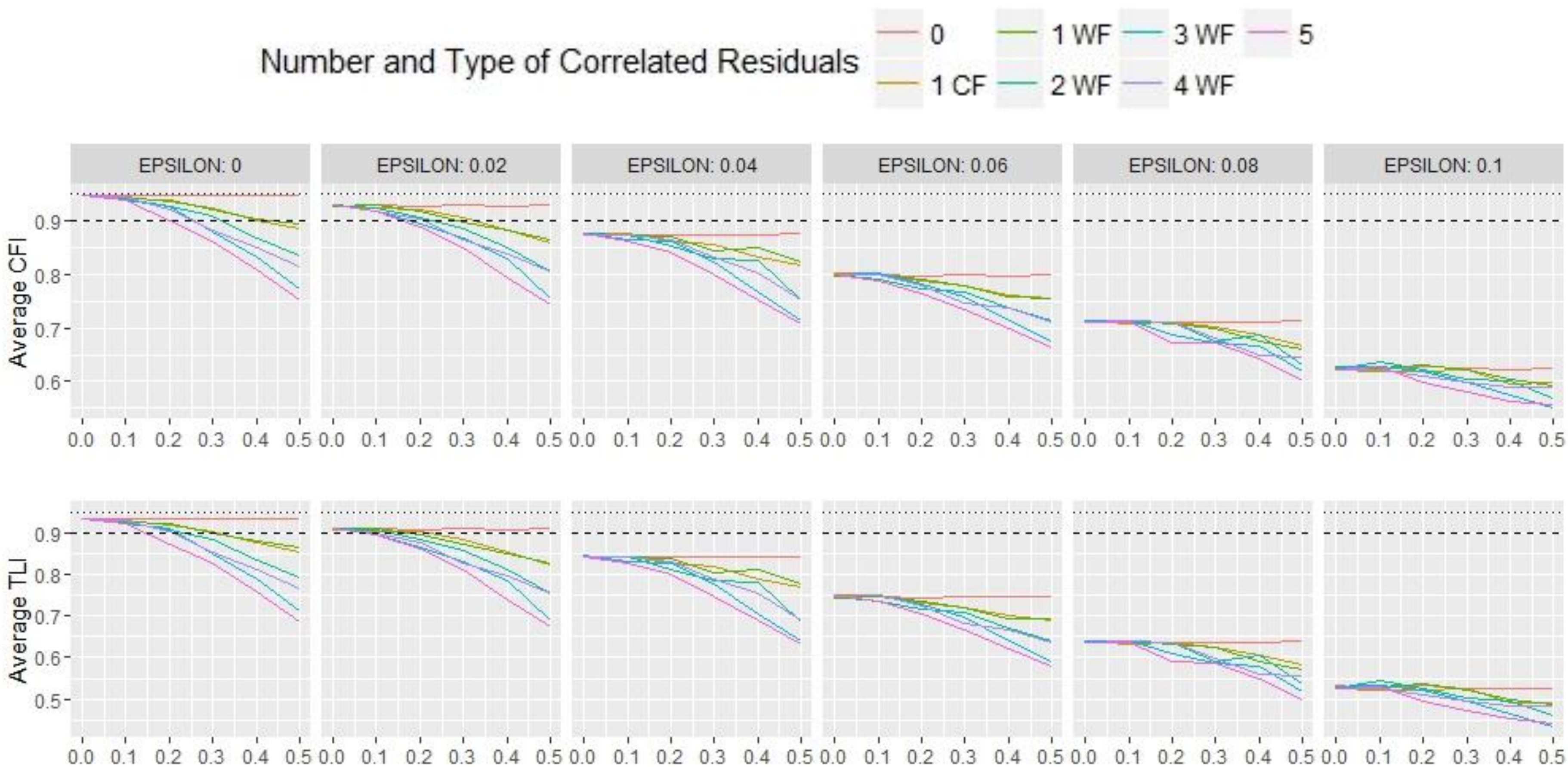
Simulation: Results SRMR



Magnitude of Correlated Residuals

ϵ = 0.2, 0.6, 0.10 omitted

Simulation: Results CFI and TLI



Magnitude of Correlated Residuals

Conclusions

Even when the model fits perfectly ($\epsilon = 0$) a few correlated residuals can result in poor fit.

As non-specific error increases, model fit gets worse

More correlated residuals and higher magnitudes of correlated residuals result in worse model fit

Within-factor and cross-factor correlated residuals impact model fit equally in the simulated case

None of the observed cutoffs would perform well at selecting the correct number of factors (in this two-factor example)

Discussion

There are many methods for selecting number of factors.

- MAP and Parallel Analysis have been shown to perform very well
- This simulation shows that model fit indices will not do well
- Convergent evidence will always reign supreme

This research provides additional evidence that blindly using cut-off values for model fit indices can be unwise.

Food for thought: Correlated residuals can be thought of as their own factor. But are they meaningful? Do we want to detect them when selecting number of factors? Are these factors **necessary**, **reliable,** and **meaningful?**

Future Directions: Examine how parallel analysis and MAP perform under these conditions.

Thank you!

National Science Foundation
WHERE DISCOVERIES BEGIN

The Mechanisms and Contingencies (MAC) Lab
The Ohio State University

Thanks to Dr. Edwards for his guidance on this project! Thanks to the Mechanisms and Contingencies Lab and the OSU Quantitative Psychology Graduate Student Meeting for your support, and to the National Science Foundation Graduate Research Fellowship, the Counsel of Graduate Students, and The Ohio State University Distinguished Dean's University Fellowship for supporting my research. And thanks to all of you for attending!

Questions?

Slides available at github.com/akmontoya/NCME2017
More information at akmontoya.com

References

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397 – 438.

Barendse, M. T., Oort, F. J., & Timmerman, M. E. (2015). Using exploratory factor analysis to determine the dimensionality of discrete responses. *Structural Equation Modeling*, *22*(1), 87-101.

Brown, T. A., White, K. S., Forsyth, J. P., & Barlow, D. H. (2004). The structure of perceived emotional control: Psychometric properties of a revised anxiety control questionnaire. *Behavior Therapy*, *35*, 75-99.

Campbell-Sills, L., Liverant, G. L., & Brown, T. A. (2004). Psychometric evaluation of the behavioral inhibition/behavior activation scales in a large sample of outpatients with anxiety and mood disorders. *Psychological Assessment*, *16*(3), 244 – 254.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J. & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods Research*, *36*(4), 462-494.

Cudeck, R. & Browne, M. W. (1992). Constructing a covariance matrix that yields a specified minimizer and a specified minimum discrepancy function value. *Psychometrika*, *57*(3), 357 – 369.

Floyd, F. J., & Widaman, K. F. (1995) Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286 – 299.

Garrido L. E., Abad, F. J., & Ponsoda, V. (2016). Are fit indices really fit to estimate the number of factors with categorical variables? Some cautionary findings via Monte Carlo simulation. *Psychological Methods*, *21*(1), 93 – 111.

Heene, M., Hilbert, S., Draxler, C., & Ziegler, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, *16*(3), 319 – 336.

MacCallum, R. C., Widaman, K. F., Preacher, K. P., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, *36*(4), 611-637.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84-99.

Marsh, H. W., Hau, K. T., & Wen, Z. (1999). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indices and dangers in overgeneralizing Hu and Bentler's (1999) finings. *Structural Equation Modeling*, *11*(3), 320 – 341.

Nye, C. D., & Drasgow, F. (2010). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods*, *14*(3), 548 – 570.

Ortuño-Sierra, J., Santarén-Rosell, M., de Albéniz, A. P., & Fonseca-Pedrero, E. (2015) Dimensional structure of the Spanish version of the positive and negative affect schedule (PANAS) in adolescents and young adults.

Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model section perspective. *Multivariate Behavioral Research*, *48*(1), 28-56.

Reichenheim, M. E., Moraes, C. L., Oliveira, S. SD., & Lobato, G. (2011). Revisitng the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): Empirical evidence for a general factor. *BMC Medical Research Methodology*, *11*(93).

Sanne, B., Torsheim, T., Heiervang, E., & Stormark, K. M. (2009). The strengths and difficulties questions in the Bergen child study: A conceptually and methodologically motivated structural analysis. *Psychological Assessment*, *21*(3), 352-364.