THE PERFORMANCE OF THE SRMR, RMSEA, CFI, AND TLI: AN EXAMINATION OF SAMPLE SIZE, PATH SIZE, AND DEGREES OF FREEDOM

Gita TAASOOBSHIRAZI[1]
Associate Professor, Department of Quantitative and Mixed Methods Research Methodologies, University of Cincinnati, USA

E-mail: gita.tshirazi@gmail.com

Shanshan WANG[2]
PhD Student, Department of Quantitative and Mixed Methods Research Methodologies, University of Cincinnati, USA

Abstract
The SRMR, RMSEA, CFI, and TLI are commonly used fit indexes reported when describing the fit of structural equation models (SEM) used in math and science education. A large number of the models tested in math and science education tend to be path models that study the interaction between various motivational, affective, contextual, and cognitive variables or latent growth curve models that examine change in students over time. The majority of these models tend to have small degrees of freedom and small sample sizes. Given the common use of these fit indexes, it is important to understand their performance when reported for relatively simple models.

Keywords: *structural equation modeling; growth model; sample size; degrees of freedom; simulation; science and math education*

The standardized root mean square residual (SRMR), root mean square error of approximation (RMSEA), comparative fit index (CFI), and the Tucker Lewis Index (TLI) are commonly used fit indexes reported when describing the fit of structural equation models (Kline, 2010; Worthington & Whittaker, 2006), with the RMSEA, SRMR, and CFI being among the most widely reported in the SEM literature (Kline, 2010).

The SRMR is a measure of the mean absolute correlation residual, with smaller values suggesting good model fit (Kline, 2010). The RMSEA provides information about 'badness of fit', with lower RMSEA values indicating good model fit (Kline, 2010). The CFI and TLI are both incremental fit indexes that assess the improvement in the fit of a model over that of a baseline model with no relationship among the model variables; larger values indicate better model fit (Kline, 2010).

Several studies have examined the performance of these and other fit indexes under various conditions (e.g., impact of sample size on fit index values) using simulations (Chen et al., 2008; Hu & Bentler, 1999). However, this work has examined the performance of the various fit indexes under different conditions with models that have moderate to large degrees of freedom. For example, (Hu & Bentler, 1999) examined the performance of the SRMR, TLI, CFI, and RMSEA with models with degrees of freedom larger than 80. There is a dearth of research examining the performance of SEM fit indexes using models with smaller degrees of freedom typical of path models or latent growth curve models tested in math and science education (Kenny, Kaniskan & McCoach, 2014; Kenny & McCoach, 2003).

Studies that have examined the impact of degrees of freedom (Breivick & Olsson, 2001; Kenny & McCoach, 2003) on model fit have tended to focus on the RMSEA and have found that the RMSEA showed better fit (smaller RMSEA values) for models with larger degrees of freedom. Kenny, Kaniskan, and McCoach (2014), who studied the performance of the RMSEA with models with small degrees of freedom, found that models with a combination of a smaller degrees of freedom and smaller sample sizes had RMSEA values that often falsely indicated a poor model fit. The authors found the RMSEA to be elevated with small sample sizes ($N \leq 100$). As the model degrees of freedom decreased, model rejection rates increased for the RMSEA, even with sample sizes as large as 1000. The RMSEA decreases if there are more degrees of freedom and/or a larger sample size, keeping everything else constant (Kline, 2010). This suggests that more parsimonious models have smaller RMSEA values. However, with the exception of the Kenny, Kaniskan, and McCoach (2014) study that looked at 1, 2, 3, 5, 10, 20, and 50 degrees of freedom with sample sizes that ranged from 50 to 1,000, other studies examining the RMSEA have used degrees of freedom much larger (e.g., 24 to 528 degrees of freedom) than those found in models frequently tested in math and science education.

The SRMR, RMSEA, CFI, and TLI are commonly used fit indexes reported when describing the fit of models used in math and science education (e.g., Bailey, Taasoobshirazi, & Carr, 2014; Byars-Winston & Fouad, 2008; Mettern & Schau, 2002; Stevens, Olivarez, Lan, & Tallent-Runnels, 2004; Glynn, Taasoobshirazi & Brickman, 2007). Many of the models tested in math and science education are path models that examine the interaction between various motivational, affective, cognitive, and contextual variables (e.g., Ha, Haury, & Nehm, 2012; Kingir, Tas, Gok, Vural, 2013; Kirbulut, 2014; Stevens et al., 2004). These models tend to have fewer than 10 variables and small degrees of freedom (less than 10) (e.g., Adedokun, Bessenbacher, Parker, & Kirkham, 2013; Akyol, Tekkaya, Sungur, & Traynor, 2012; Bailey et al., 2014). In addition, a common structural equation model with a small degrees of freedom tested in math and science education is the latent growth model (e.g., Gottfried, Marcoulides, Gottfried, & Oliver, 2009). Given the common use of these fit indexes, it is important to understand their performance when used with models with small degrees of freedom.

Also of interest was the impact of sample size on the performance of the fit indexes with models with small degrees of freedom. For example, the research in math and science education testing path models often have sample sizes that tend to be less than $N = 250$, with many studies having sample sizes less than $N = 100$ (e.g., Bailey et al., 2014; Ha, Haury, & Nehm, 2012). This is much smaller than the sample sizes typically found in simulation studies on fit indexes in SEM (e.g., Hu & Bentler, 1999). Several studies have shown the importance of sample size on the performance of fit indexes such as the RMSEA (Chen et al., 2008; Kenny et al., 2014). Specifically, Type II error rates for the fit indexes increase as sample size decreases. A goal of the present study was to determine the interaction between sample size and degrees of freedom on the performance of the SRMR, RMSEA, CFI, and TLI for models that have small degrees of freedom and small sample sizes, typical of what is found in the path models and latent growth curve models tested in math and science education.

We wanted to know, for our four fit indexes and when working with models with small degrees of freedom: What is the performance of these fit indexes and their rejection rates across various sample size and degrees of freedom combinations? Specifically, do models with smaller degrees of freedom (more paths in the model) require a larger sample size, similar to the results of the Kenny, Kaniskan, and McCoach (2014) findings for the RMSEA or is a smaller sample size sufficient for models that have smaller degrees of freedom in line with research illustrating that adding paths to a model tends to improve fit (Kline, 2010)? What is a sufficient sample size for a small degrees of freedom model needed to avoid making a type II error?

Method

We conducted a Monte Carlo simulation to test correctly specified growth models with varying sample sizes and degrees of freedom. To do so, we followed the same simulation techniques as Kenny et al. (2014). Specifically, "intercept loadings were all fixed to one and slope loadings were fixed to zero for wave 1 and increased in one-unit increments thereafter. The population mean of the intercept factor was 0.5 and the variance was set at 1.0: The population mean of the slope factor was 1.0 and its variance was 0.2. The covariance between slope and intercept was 0.1, and all error variances were set at 0.5. The models were as follows based and are designated by their degrees of freedom: df = 1: 3-wave growth model, df = 2: 3-wave growth model with equal error variances and the loading for the slope factor at wave 3 free, df = 3: 3-wave growth model, with equal error variances, df = 5: 4-wave growth model, df = 10: 5 wave growth model (figure 1), df = 20: 7 wave growth model, with loadings on the slope factor for the last three times free, and df = 50: 10 wave growth model" (Kenny et al., 2014, p. 10). In addition, we used the same sample sizes as Kenny et al., (2014) including: 50, 100, 200, 400, 600, and 1,000.

The simulation feature of Mplus Version 7.3 was used to generate the data, and the R package MplusAutomation (Hallquist & Wiley, 2014) was used to estimate the parameters for each simulation condition. We replicated each condition 1,000 times. The raw data were generated from a multivariate normal distribution. In order to address our research questions, we calculated the average value of the model fit indices for each simulation condition and their accompanying rejection rates defined as the number of times that the model fit indices exceeded recommended cutoff values specified in the research. Specifically, a CLI cutoff value of .95, a TLI cutoff value of .95, a RMSEA cutoff value of .05, and a SRMR cutoff value of .08 were used (Hooper, Coughlan, & Mullen, 2008).

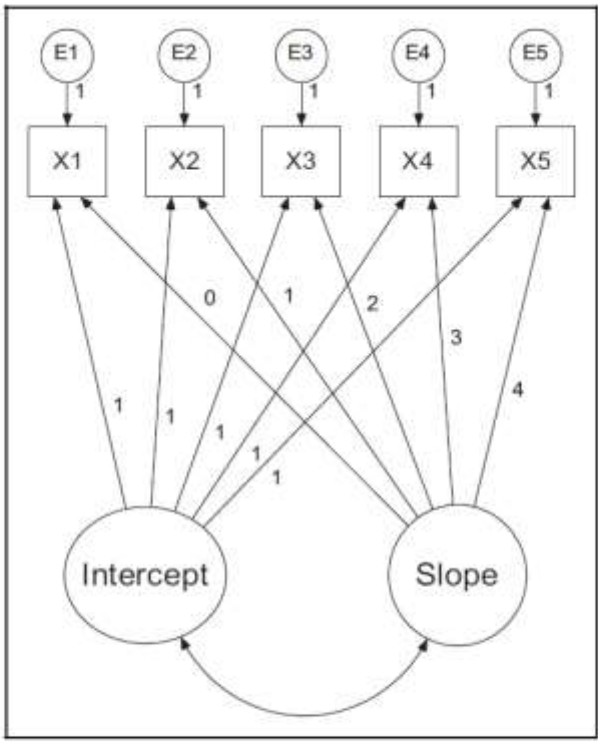

Figure 1. Simulation model for 10 degrees of freedom.

Results

CFI. Results of the simulation for the various degrees of freedom and sample sizes are presented in Table 1. The table reports CFI values and accompanying rejection rates. Figure 2 shows the CFI values for the selected sample sizes and degrees of freedom. Results indicated that adding paths to the model (decreasing the degrees of freedom) did not do much in terms of altering CFI values. Increasing sample size resulted in larger CFI values. Based on rejection rates, we almost never found the CFI to fall below the cutoff for sample sizes 100 or larger. For N = 50, the largest rejection rate was about 10%, suggesting that such small sample sizes may be problematic when using the CFI (unless degrees of freedom are very large).

Table 1. CFI values for selected degrees of freedom and sample size.
Rejection rates are in parentheses

df	1	2	3	5	10	20	50
N = 50	0.992 (3.90%)	0.990 (6.00%)	0.986 (9.80%)	0.990 (5.10%)	0.991 (2.30%)	0.989 (3.20%)	0.992 (0.10%)
N = 100	0.996 (0.30%)	0.995 (1.60%)	0.993 (2.20%)	0.995 (0.20%)	0.996 (0.10%)	0.995 (0.10%)	0.997 (0.00%)
N = 200	0.998 (0.00%)	0.997 (0.20%)	0.996 (0.20%)	0.997 (0.00%)	0.998 (0.00%)	0.998 (0.00%)	0.999 (0.00%)
N = 400	0.999 (0.00%)	0.999 (0.00%)	0.998 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)
N = 600	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	1.000 (0.00%)
N = 1000	1.000 (0.00%)	1.000 (0.00%)	0.999 (0.00%)	0.999 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)

TLI. Results of the simulation for the various degrees of freedom and sample sizes are presented in Table 2.

Table 2. TLI values for selected degrees of freedom and sample size.
Rejection rates are in parentheses

df	1	2	3	5	10	20	50
N = 50	0.997 (16.10%)	1.001 (12.20%)	0.998 (9.80%)	0.998 (7.30%)	0.998 (2.30%)	0.995 (3.70%)	0.995 (0.10%)
N = 100	0.999 (8.30%)	0.999 (4.00%)	0.999 (2.20%)	0.999 (0.90%)	0.999 (0.10%)	0.999 (0.10%)	0.999 (0.00%)
N = 200	0.999 (2.10%)	0.999 (0.40%)	0.999 (0.20%)	0.999 (0.00%)	1.000 (0.00%)	0.999 (0.00%)	1.000 (0.00%)
N = 400	1.000 (0.20%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)
N = 600	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)
N = 1000	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)	1.000 (0.00%)

The table reports TLI values and accompanying rejection rates. Figure 3 shows the TLI values for the selected sample sizes and degrees of freedom. Results indicated that adding paths to the model (decreasing the degrees of freedom) did not do much in terms of altering TLI values. Increasing sample size resulted in larger TLI values. The model rejection rate did increase to 16% for a model with one degrees of freedom and a sample size of 50, suggesting that we could reject a correctly specified model when using the TLI and a combination of a small sample size and such a small degrees of freedom.

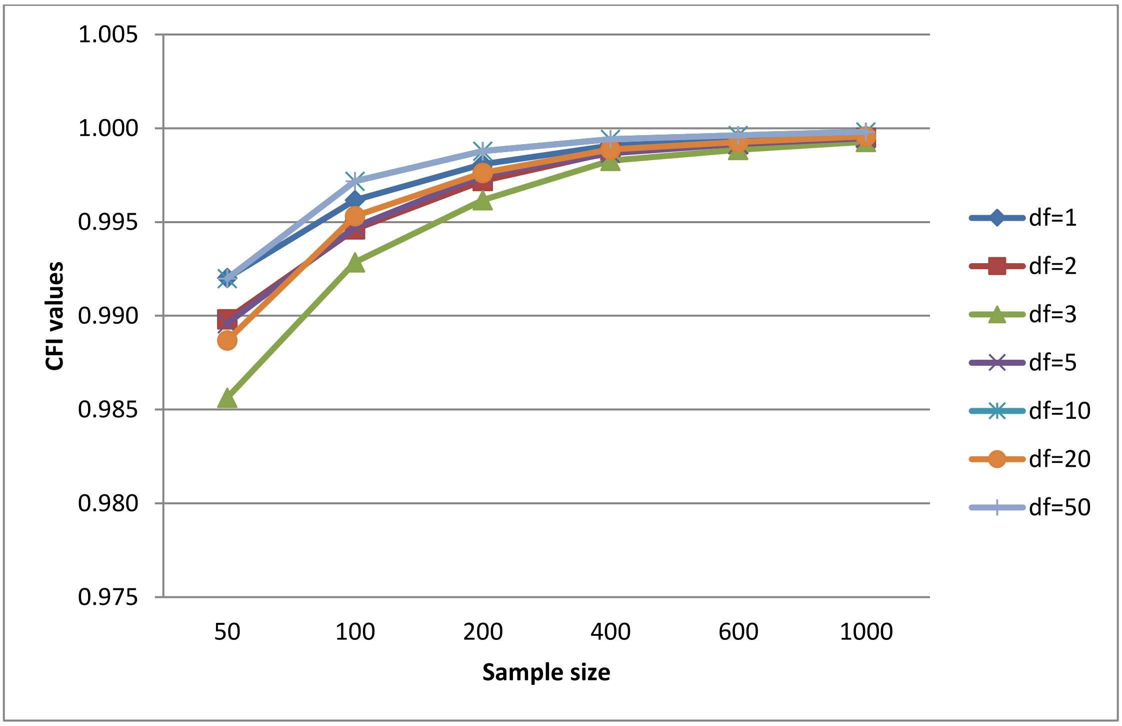

Figure 2. CFI values for selected degrees of freedom and sample size

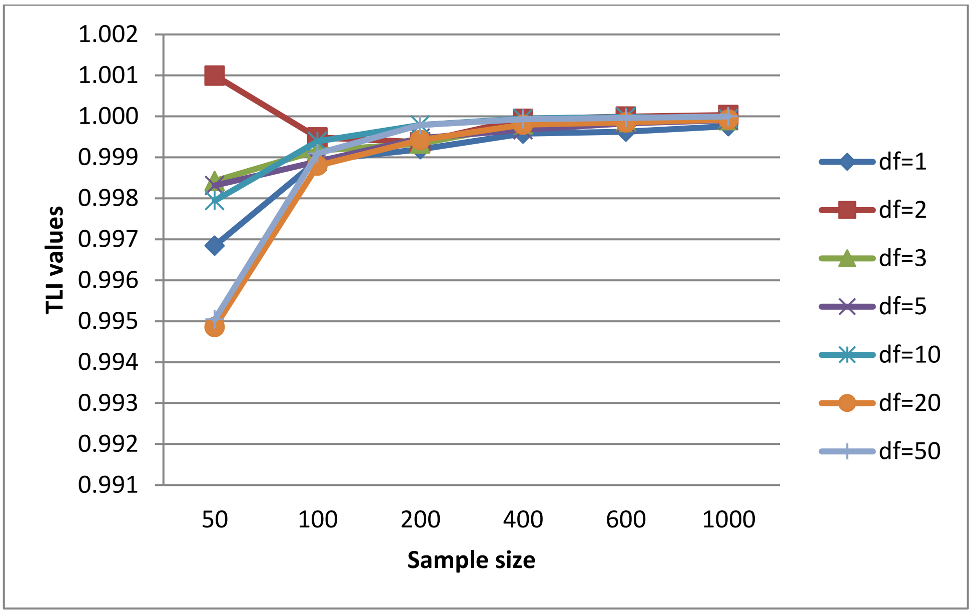

Figure 3. TLI values for selected degrees of freedom and sample size.

RMSEA. Results of the simulation for the various degrees of freedom and sample sizes are presented in Table 3. The table reports RMSEA values and accompanying rejection rates. Figure 4 shows the RMSEA values for the selected sample sizes and degrees of freedom.

Table 3. RMSEA values for selected degrees of freedom and sample size. Rejection rates are in parentheses.

Df	1	2	3	5	10	20	50
N = 50	0.053 (29.50%)	0.045 (30.10%)	0.048 (34.50%)	0.044 (34.20%)	0.041 (34.80%)	0.041 (35.60%)	0.041 (37.00%)
N = 100	0.037 (25.40%)	0.034 (26.30%)	0.034 (28.70%)	0.032 (26.60%)	0.027 (21.00%)	0.025 (16.00%)	0.021 (8.10%)
N = 200	0.027 (21.10%)	0.025 (20.10%)	0.025 (18.80%)	0.022 (14.60%)	0.019 (8.40%)	0.018 (2.70%)	0.014 (0.30%)
N = 400	0.019 (14.20%)	0.017 (8.00%)	0.016 (6.50%)	0.016 (4.30%)	0.013 (1.00%)	0.012 (0.00%)	0.009 (0.00%)
N = 600	0.016 (8.40%)	0.013 (5.00%)	0.014 (2.60%)	0.013 (0.90%)	0.010 (0.00%)	0.010 (0.00%)	0.008 (0.00%)
N = 1000	0.013 (4.90%)	0.010 (1.60%)	0.011 (0.70%)	0.010 (0.20%)	0.008 (0.00%)	0.007 (0.00%)	0.006 (0.00%)

Results indicated that adding paths to the model (decreasing the degrees of freedom) tend to result in larger RMSEA values. Increasing sample size resulted in smaller RMSEA values. It is also important to note that model rejection rates were high (greater than 30%) with a sample size of 50 regardless of the degrees of freedom. In addition, for an N of 100 and 200, models with smaller degrees of freedom had higher rejection rates. For example, the RMSEA exceeded the cutoff of .05 nearly 29% of the time with a sample size of 100 and degrees of freedom of 3. This is in line with findings by Kenny et al. (2014) indicating that researchers should proceed with caution when using the RMSEA with SEM models with small degrees of freedom and small sample sizes.

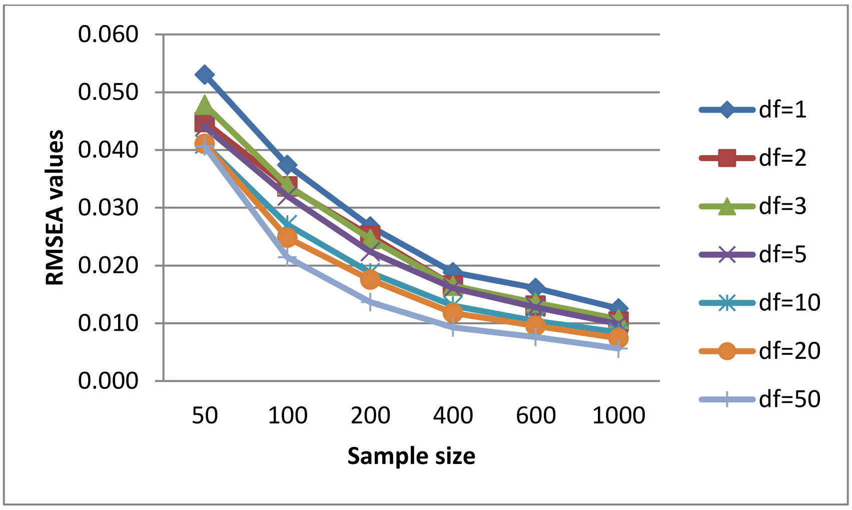

Figure 4. RMSEA values for selected degrees of freedom and sample size

SRMR. Results of the simulation for the various degrees of freedom and sample sizes are presented in Table 4. The table reports SRMR values and accompanying rejection rates. Figure 5 shows the SRMR values for the selected sample sizes and degrees of freedom.

Table 4. SRMR values for selected degrees of freedom and sample size.
Rejection rates are in parentheses.

df	1	2	3	5	10	20	50
N = 50	0.022 (0.20%)	0.047 (13.60%)	0.056 (18.50%)	0.053 (12.40%)	0.059 (15.80%)	0.064 (18.70%)	0.052 (3.10%)
N = 100	0.015 (0.00%)	0.034 (3.90%)	0.039 (4.30%)	0.037 (1.40%)	0.040 (1.20%)	0.043 (0.40%)	0.035 (0.00%)
N = 200	0.010 (0.00%)	0.024 (0.50%)	0.027 (0.20%)	0.026 (0.00%)	0.028 (0.00%)	0.031 (0.00%)	0.024 (0.00%)
N = 400	0.007 (0.00%)	0.017 (0.00%)	0.019 (0.00%)	0.018 (0.00%)	0.020 (0.00%)	0.022 (0.00%)	0.017 (0.00%)
N = 600	0.006 (0.00%)	0.013 (0.00%)	0.015 (0.00%)	0.015 (0.00%)	0.016 (0.00%)	0.018 (0.00%)	0.014 (0.00%)
N = 1000	0.005 (0.00%)	0.010 (0.00%)	0.012 (0.00%)	0.011 (0.00%)	0.013 (0.00%)	0.014 (0.00%)	0.011 (0.00%)

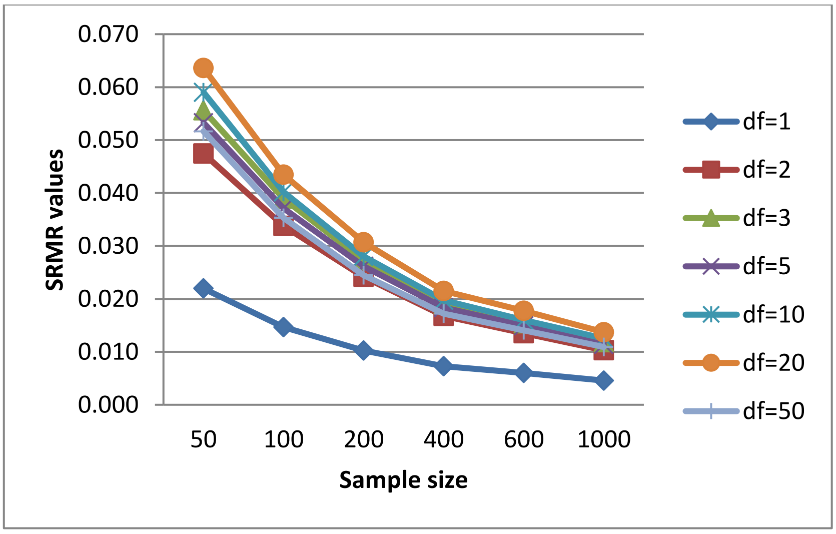

Figure 5. SRMR values for selected degrees of freedom and sample size.

Results indicated that adding paths to the model (decreasing the degrees of freedom) tend to result in smaller SRMR values. Increasing sample size resulted in smaller SRMR values. Once again, rejection rates tend to increase for a smaller sample size of 50 regardless of the degrees of freedom.

Conclusion

This is one of the first studies to examine the performance of the SRMR, RMSEA, CFI, and TLI with models with the small degrees of freedom typical of models found in math and science education. Kenny, Kaniskan & McCoach (2014) examined the RMSEA for correctly specified growth models with small degrees of freedom and we hoped to extend this line of research by examining the performance the RMSEA and additional fit indexes. We chose our four fit indexes because they are among the most widely reported in the SEM literature (Kline, 2010). Because we tested correctly specified models, we hoped that our fit indexes would not violate the cutoffs reported in the research and that rejection rates would be low. We manipulated sample size across various and small degrees of freedom and found that, in general, researchers should avoid sample sizes less than 100 when testing small degrees of freedom models. In fact, science and math education researchers should avoid reporting the RMSEA when sample sizes are smaller than 200, particularly when combined with small degrees of freedom. Small degrees of freedom do not tend to result in rejection of correctly specified models for the TLI, CFI, and SRMR, particularly if they tested using larger sample sizes.

References

1. Adedokun, O. A., Bessenbacher, A. B., Parker, L. C. Kirkham, L. L. and Burgess, W. D. **Research skills and STEM undergraduate research students' aspirations for research careers: Mediating effects of research self-efficacy,** Journal of Research in Science teaching, Vol. 50, No. 8, 2013, pp. 940-951
2. Akyol, G., Tekkaya, C., Sungur, S. and Traynor, A. **Modeling the interrelationships among pre-service science teachers' understanding and acceptance of evolution, their views on nature of science and self-efficacy beliefs regarding teaching evolution,** Journal of Science Teacher Education, Vol. 23, No. 8, 2012, pp. 937-957
3. Bailey, M., Taasoobshirazi, G. and Carr, M. **A Multivariate Model of Achievement in Geometry,** The Journal of Educational Research, Vol. 107, No. 6, 2014, pp. 440-461
4. Breivik, E., and Olsson, U. H. **Adding variables to improve fit: The effect of model size on fit assessment in LISREL,** Structural equation modeling: Present and future, 2001, pp. 169-194
5. Byars-Winston, A. M. and Fouad, N. A. **Math and science social cognitive variables in college students contributions of contextual factors in predicting goals,** Journal of Career Assessment, Vol. 16, No. 4, 2008, pp. 425-440
6. Chen, F., Curran, P. J. Bollen, K. A., Kirby, J. and Paxton, P. **An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models,** Sociological methods & research, Vol. 36, No. 4, 2008, pp. 462-494

7. Glynn, S. M., Taasoobshirazi, G. and Brickman, P. **Nonscience majors learning science: A theoretical model of motivation.** Journal of Research in Science Teaching, Vol. 44, No. 8, pp. 2007, pp. 1088-1107
8. Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W. and Oliver, P. H. **A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation.** Journal of Educational Psychology, Vol. 101, No. 3, 2009, p. 729
9. Ha, M., Haury, D. L. and Nehm, R. H. **Feeling of certainty: uncovering a missing link between knowledge and acceptance of evolution,** Journal of Research in Science Teaching, Vol. 49, No. 1, 2012, pp. 95-121
10. Hallquist, M. and Wiley, J. **Package 'MplusAutomation'.** 2013, retrieved from http://CRAN.Rproject.org/package=MplusAutomation
11. Hooper, D., Coughlan, J. and Mullen, M. **Structural equation modelling: Guidelines for determining model fit.** Articles, 2008, p. 2
12. Hu, L.-T. and Bentler, P. M. **Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.** Structural equation modeling: a multidisciplinary journal, Vol. 6, No. 1, 1999, pp. 1-55
13. Kenny, D. A., Kaniskan, B. and McCoach, D. B. **The performance of RMSEA in models with small degrees of freedom.** Sociological Methods & Research, 2014, 0049124114543236.
14. Kenny, D. A. and McCoach, D. B. **Effect of the number of variables on measures of fit in structural equation modeling.** Structural equation modelling, Vol. 10, No. 3, 2003, pp. 333-351
15. Kirbulut, Z. D. **Modeling the Relationship between High School Students' Chemistry Self-Efficacy and Metacognitive Awareness.** International Journal of Environmental and Science Education, Vol. 9, No. 2, 2014, pp. 177-196
16. Kingir, S., Tas, Y., Gok, G. and Vural, S. S. **Relationships among constructivist learning environment perceptions, motivational beliefs, self-regulation and science achievement.** Research in Science & Technological Education, Vol. 31, No. 3, 2013, pp. 205-226
17. Kline, R. B. **Principles and Practice of Structural Equation Modeling, 3rd edn Guilford Press.** New York, USA, 2010
18. Mattern, N., and Schau, C. **Gender differences in science attitude-achievement relationships over time among white middle-school students.** Journal of Research in Science Teaching, Vol. 39, No. 4, 2002, pp. 324-340
19. Muthén, B. O. **Mplus technical appendices.** Los Angeles, CA: Muthén & Muthén,1998
20. Stevens, T., Olivarez, A., Lan, W. Y. and Tallent-Runnels, M. K. **Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity.** The Journal of Educational Research, Vol. 97, No. 4, 2004, pp. 208-222
21. Worthington, R. L., and Whittaker, T. A. **Scale development research a content analysis and recommendations for best practices.** The Counseling Psychologist, Vol. 34, No. 6, 2006, pp. 806-838

[1]Gita Taasoobshirazi is an Associate Professor of Quantitative and Mixed Methods Research Methodologies at the University of Cincinnati. Her work involves examining and advancing the use of quantitative methods in education.

[2]Shanshan Wang is a doctoral student in Quantitative and Mixed Methods Research Methodologies at the University of Cincinnati. Her work involves understanding and advancing quantitative research methods in the social sciences.

HEURISTIC DECISION MAKING UTILIZING COMPLETE TOURNAMENTS[1]

Johan KOK[2]
Business Community Safety & Development Department
City of Tshwane, Republic of South Africa

E-mail: kokkiek2@tshwane.gov.za

R. JAYASREE
Research Scholar, Department of Mathematics,
Nirmala College for Women Coimbatore, India

E-mail: sree.ramachandran14@gmail.com

Abstract

The paper presents the newly developed Abarnica heuristic ranking method applied to fuzzy-bias decision making with the central technique derived from complete tournaments or derivatives thereof. An useful recursive proposition called, Akhil's proposition together with a challenging conjecture is presented. The challenge to derive an efficient algorithm to apply the Abarnica heuristic method which appears to be very efficient with manual applications remains open. The authors advocate that the main advantage of the Abarnica heuristic ranking method is that strong bias are analytically mitigated in fuzzy-bias decision making applications which require ranking.

Keywords: *Complete tournament, Perron-Frobenius theorem, Abarnica heuristic, corrupt driving testing*

1. Introduction

For general notation and concepts in graphs and digraphs see [1, 2, 3]. Unless mentioned otherwise all graphs are simple, connected and directed graphs (digraphs). Furthermore, if the context is clear the terms *vertex* and *player* will be used interchangeably.

A directed complete graph of order $n \geq 1$ with vertex set $V(G) = \{v_1, v_2, ..., v_n\}$ can be considered to be complete tournament, denoted by T_n. The understanding of a complete tournament is that all distinct pairs of vertices are players in a match and the arc (v_i, v_j)