Structural Equation Modeling: A Multidisciplinary Journal, 00: 1–6, 2016

ISSN: 1070-5511 print / 1532-8007 online
DOI: 10.1080/10705511.2016.1225260

TEACHER'S CORNER

New Ways to Evaluate Goodness of Fit: A Note on Using Equivalence Testing to Assess Structural Equation Models

Katerina M. Marcoulides[1] and Ke-Hai Yuan[2]
[1]*Arizona State University*
[2]*University of Notre Dame*

Structural equation models are typically evaluated on the basis of goodness-of-fit indexes. Despite their popularity, agreeing what value these indexes should attain to confidently decide between the acceptance and rejection of a model has been greatly debated. A recently proposed approach by means of equivalence testing has been recommended as a superior way to evaluate the goodness of fit of models. The approach has also been proposed as providing a necessary vehicle that can be used to advance the inferential nature of structural equation modeling as a confirmatory tool. The purpose of this article is to introduce readers to key ideas in equivalence testing and illustrate its use for conducting model–data fit assessments. Two confirmatory factor analysis models in which a priori specified latent variable models with known structure and tested against data are used as examples. It is advocated that whenever the goodness of fit of a model is to be assessed researchers should always examine the resulting values obtained via the equivalence testing approach.

Keywords: CFI, equivalence testing, fit indexes, likelihood ratio statistic, RMSEA

Structural equation modeling (SEM) methods are extremely popular in many scientific disciplines because they enable researchers to identify and test models that can account for complex multivariate relationships among observed and latent variables. Evaluating the fit of a proposed model is typically carried out on the basis of so-called goodness-of-fit indexes. These can include an inferential goodness-of-fit index (a chi-square value) as well as a number of other descriptive or alternative fit indexes (e.g., the root mean square error of approximation [RMSEA], the comparative fit index [CFI]; Bentler, 1990; Steiger & Lind, 1980). Most currently available SEM computer programs (e.g., AMOS, EQS, LISREL, Lavaan, M*plus*, OpenMx) automatically supply dozens of these fit indexes as standard output.[1] It has been suggested that to support model fit based on these indexes, the following are needed: a nonsignificant χ^2 goodness-of-fit value, a CFI > .90, and an RMSEA below .05 with the left endpoint of its 95% confidence interval (CI) including 0 (Browne & Cudeck, 1992; Raykov & Marcoulides, 2006).

There has been much debate on the use of goodness-of-fit indexes, particularly on what value these indexes should attain for one to confidently decide between the acceptance and rejection of a model (usually referred to as cutoff values or criteria). For example, cutoff values for both the RMSEA (0.01, 0.05, 0.08, and 0.10) and the CFI (0.99, 0.95, 0.92, and 0.90) have been commonly used to distinguish between

Correspondence should be addressed to Katerina M. Marcoulides, Department of Psychology, Arizona State University, P.O. Box 871104, Tempe, AZ 85287. E-mail: kmarcoul@asu.edu

[1] For a detailed discussion on these computer programs, we refer readers to their manuals and various related publications (e.g., Arbuckle, 2012; Bentler, 2006; Boker et al., 2011; Byrne, 2001; Jöreskog & Sörbom, 1993; Muthén & Muthén, 2014; Pritikin, Hunter, & Boker, 2015; R Development Core Team, 2011; Raykov & Marcoulides, 2006; Rosseel, 2012).

excellent, close, fair, and mediocre or poor models, respectively. Because fit is recognized to be a matter of degree, Hu and Bentler (1998, 1999) provided some general guidelines concerning the use of cutoff criteria for these various fit indexes. Unfortunately, and despite the countless cautions offered by Hu and Bentler (1998, 1999) and other researchers (e.g., Marsh, Balla, & Hau, 1996; Marsh, Hau, & Wen, 2004) about indiscriminately using these cutoff criteria, many researchers have inaccurately interpreted the recommended guidelines, with some going so far as to suggest they be arbitrarily increased or even completely abandoned (Barrett, 2007; Russell, 2002). For example, Barrett (2007) declared, "I would recommend banning ALL such indices from ever appearing in any paper as indicative of 'model acceptability' or 'degree of misfit'" (p. 821). However, many other researchers believe that there can still be a place for model fit indexes, as long as no decision concerning goodness of fit is based on just a single index, no matter how favorable for the model that index might appear (Hu & Bentler, 1999; Raykov & Marcoulides, 2006, 2008; Tomarken & Waller, 2005; Yu, 2002). Considering the complexity of issues related to the potential sensitivity of these indexes to different types of modeling and data conditions, these researchers have also suggested that each fit index should be seen as merely representing a specific aspect of the fit of a proposed model.

Assessing model fit in SEM is further complicated by the fact that, even when a model is not rejected based on these indexes, a researcher still cannot claim that the model is correctly specified or that the quality of the model is good. For example, a χ^2 goodness-of-fit value represents a test statistic of the fit of a model when testing the null hypothesis that the proposed model fits the corresponding population covariance matrix perfectly. A nonsignificant χ^2 goodness-of-fit value does not enable a researcher to claim that the model is correctly specified or that the quality of the model is good. Getting a nonsignificant result only implies that there is not enough evidence to reject the null hypothesis that the proposed model fits the corresponding population covariance matrix perfectly.

Realizing the limited utility of this approach for evaluating model fit in SEM, Yuan, Chan, Marcoulides, and Bentler (2016) recently proposed the use of equivalence testing with adjusted fit indexes. They also advocated that all researchers should start reporting equivalence testing results to convey the goodness of fit of SEM models.[2] It has been proposed that the equivalence testing approach can also provide a necessary vehicle that can be used to advance the inferential nature of SEM as a confirmatory tool. According to Yuan et al. (2016), "equivalence testing gives SEM the needed property to be a scientific methodology" (p. 327). In contrast to the current state of affairs in SEM, where researchers simply report fit index values meeting certain criteria to support a proposed model, equivalence testing provides additional information about the desired confidence for the proposed model.

To date, equivalence testing has rarely been used in SEM applications; this article is intended to introduce readers to key ideas illustrating how it can be used to conduct model–data fit assessments in SEM. A more detailed treatment of equivalence testing in SEM complete with mathematical proofs can be found in Yuan et al. (2016). For a general treatment of the topic in statistical hypothesis testing, readers are referred to Wellek (2010). Yuan et al. (2016) also contrasted the similarities and differences of equivalence testing with traditional hypothesis-testing-based approaches to model fit. We do not attempt to repeat these aspects about equivalence testing, but rather highlight the potential benefits and research implications that can arise when using this approach. The new approach is illustrated using two simple confirmatory factor analysis (CFA) examples in which a priori specified latent variable models with known structure are tested against data. As the illustrated fit assessment method is identical irrespective of the complexity of the examined model, the examples are purposefully kept simple to emphasize key ideas, benefits, and implications. The central issue is the overall assessment of model fit based on examining both derived conventional criteria and equivalence testing criteria.[3] The examination of model fit criteria is a widespread practice in similar circumstances in present-day empirical research using structural equation models. It is important, therefore, that researchers appreciate the tremendous benefits of equivalence testing in practical terms and use it whenever SEM is selected as the relevant methodology. All model fit results reported in this article were obtained using the widely circulated freeware program R and the commercially available computer software M*plus* (although any other available SEM computer program could be used to obtain the same results).

CONDUCTING EQUIVALENCE TESTING

To conduct equivalence testing in SEM, the so-called T-size (minimum tolerable size) of misspecification (ϵ_t) corresponding to the observed likelihood ratio test statistic (T_{ML}) based on maximum likelihood estimation is needed.[4] Although this T-size of misspecification can be related to any fit index, to keep with the approach presented in Yuan et al. (2016), we

[2] Although Yuan et al. (2016) focused mainly on using equivalence testing with SEM models, it was indicated that the same approach could easily be applied to all types of data analysis model (time series, item response models, etc.).

[3] Although equivalence testing of an overall model structure is only illustrated in this article, such an approach could also be used to examine individual aspects of a model.

[4] Although the statistic is only examined here with normally distributed data, as indicated by Yuan et al. (2016), with other distributions perhaps robust transformations of the observed data or normal distribution approximations can be effectively applied.

relate it to only the CFI and the RMSEA fit indexes. Accordingly, their corresponding T-size fit indexes are denoted, respectively, as CFI_t and $RMSEA_t$. A readily available R program that computes both these T-size indexes is available at http://www3.nd.edu/~kyuan/EquivalenceTesting/T-size_RMSEA_CFI.R. For any model being examined, the program only requires as input the obtained likelihood ratio test statistic (T_{ML}) and its degrees of freedom (df), the likelihood ratio statistic corresponding to the independence model ($T_{ML\text{-}Independence}$), the sample size ($N$), the number of observed variables (p), and the desired significance level (α). Because the degrees of freedom for the likelihood ratio statistic corresponding to the independence model can be computed using the inputted number of observed variables, this value does not need to be manually entered. Once these model details are entered into the current R session, users can then run the program. For example, when examining a model with $T_{ML} = 50$, $df = 30$, $T_{ML\text{-}Independence} = 1{,}000$, $N = 500$, $p = 20$, and $\alpha = .05$, the following Input commands would be specified:

```
#———Input————————#;
N=500; p=20;
T_ml= 50; df=30;
T_mli= 1000;
alpha=.05;
df_i=p*(p+1)/2-p;
#————————————————#;
```

The returned output will include the conventional values for both the CFI and the RMSEA, along with the equivalence testing based T-size fit indexes CFI_t and $RMSEA_t$. In addition to these T-size values, adjusted cutoff values for each T-size fit index can also be obtained using R programs available at http://www3.nd.edu/~kyuan/EquivalenceTesting/CFI_e.R and at http://www3.nd.edu/~kyuan/EquivalenceTesting/RMSEA_e.R. To use these additional programs the only inputs needed are the degrees of freedom (df) and sample size (N).

ILLUSTRATIVE EXAMPLE

For the purpose of demonstrating the equivalence testing method in this article, two simple CFA examples in which a priori specified latent variable models with known structure are tested against data—one in which a one-factor model is tested and the second in which a two-factor model is tested. In both cases, the data set was simulated based on $N = 600$ cases from $k = 2$ factors each with seven and five observed variables, respectively, which were generated according to the following model:

$$
\begin{aligned}
y_1 &= 3 + .6\eta_1 + \varepsilon_1,\\
y_2 &= 3 + .65\eta_1 + \varepsilon_2,\\
y_3 &= 3 + .7\eta_1 + \varepsilon_3,\\
y_4 &= 3 + .75\eta_1 + \varepsilon_4,\\
y_5 &= 3 + .8\eta_1 + \varepsilon_5,\\
y_6 &= 3 + .6\eta_1 + \varepsilon_6,\\
y_7 &= 3 + .65\eta_1 + \varepsilon_7,\\
y_8 &= 3 + .8\eta_2 + \varepsilon_8,\\
y_9 &= 3 + .8\eta_2 + \varepsilon_9,\\
y_{10} &= 3 + .8\eta_2 + \varepsilon_{10},\\
y_{11} &= 3 + .8\eta_2 + \varepsilon_{11}, \text{and}\\
y_{12} &= 3 + .8\eta_2 + \varepsilon_{12},
\end{aligned}
\tag{1}
$$

where η_1 and η_2 were standard normal variates with correlation $\rho = .75$, and the error terms were independent standard normal variables (the data simulation process was based on that in Raykov, Marcoulides, & Tong, 2015). The resulting covariance matrix among the $p = 12$ observed variables is provided in Table 1.

First, a one-factor model was fit to all $p = 12$ observed variables using the M*plus* program (based on maximum likelihood estimation; e.g., Bollen, 1989). Based on this model, the following conventional goodness-of-fit indexes were obtained: $\chi^2(54) = 140.089$, CFI = 0.941, and RMSEA = .052, with a 90% CI of [.041, .062]. The complete model fit information provided by M*plus* is presented in Table 2. As can been seen by examining these results, model fit information can be considered indicative of a tenable model that achieves close fit. Based on an examination of the conventional fit criteria, it can therefore be concluded that these findings suggest the one-factor model as a plausible model for the analyzed data set.

Next, entering the appropriate values as input into the previously mentioned R program to determine the equivalence testing T-size indexes, the following output is provided:

```
#————For T-size RMSEA————#;
Conventional RMSEA = 0.05158976
T-size RMSEA in equivalence testing = 0.06215945
#—————For T-size CFI—————#;
Conventional CFI = 0.9414483
T-size CFI in equivalence testing = 0.8995516
#———————————————————————#;
```

These results indicate that we are 95% confident that the population CFI is above 0.8995516 and that the size of misspecification is no more than 0.06215945, as measured by the T-size RMSEA. Evaluating next these T-size results relative to excellent, close, fair, and mediocre or poor cutoff values as determined by the appropriate R program, we obtain the following output:

TABLE 1
Covariance Matrix for the p = 12 Observed Variables

Variable	*1.*	*2.*	*3.*	*4.*	*5.*	*6.*	*7.*	*8.*	*9.*	*10.*	*11.*	*12.*
1.	1.321											
2.	0.443	1.410										
3.	0.283	0.507	1.485									
4.	0.379	0.526	0.542	1.547								
5.	0.462	0.466	0.411	0.527	1.524							
6.	0.316	0.392	0.370	0.418	0.496	1.441						
7.	0.392	0.404	0.352	0.481	0.478	0.387	1.422					
8.	0.404	0.342	0.389	0.449	0.426	0.391	0.405	1.566				
9.	0.398	0.493	0.437	0.450	0.447	0.480	0.412	0.657	1.646			
10.	0.313	0.423	0.372	0.379	0.398	0.300	0.351	0.538	0.599	1.675		
11.	0.374	0.448	0.359	0.368	0.348	0.404	0.335	0.591	0.608	0.659	1.630	
12.	0.381	0.486	0.387	0.359	0.370	0.438	0.371	0.556	0.690	0.529	0.640	1.673

Note. 1–12 = observed variables.

TABLE 2
M*plus* Model Fit Information for One-Factor Model

Number of free parameters	36
Log-likelihood	
H0 value	−11035.710
H1 value	−10965.666
Information criteria	
Akaike (AIC)	22143.420
Bayesian (BIC)	22301.710
Sample-size-adjusted BIC	22187.420
(n^* = (n + 2) /24)	
Chi-square test of model fit	
Value	140.089
Degrees of freedom	54
p value	0.0000
Root mean square error of approximation (RMSEA)	
Estimate	0.052
90% CI	[0.041, 0.062]
Probability RMSEA ≤ .05	0.387
CFI/TLI	
CFI	0.941
TLI	0.928
Chi-square test of model fit for the baseline model	
Value	1536.307
Degrees of freedom	66
p value	0.0000
Standardized root mean square residual (SRMR)	
Value	0.039

Note. CFI = comparative fit index; TLI = Tucker–Lewis Index.

```
#————————————————————————————#;
CFI_e90      CFI_e92      CFI_e95      CFI_e99
[1,] 0.8665537   0.8904456    0.9270826   0.9776777
–poor– 0.867 –mediocre– 0.891 –fair– 0.927 –close–
0.978 –excellent—
RMSEA_e01  RMSEA_e05RMSEA_e08 RMSEA_e10
[1,] 0.02828466   0.06034854   0.08933914  0.1090103
–excellent– 0.028 –close– 0.060 –fair– 0.089 –mediocre–
0.109 –poor–
#————————————————————————————#;
```

The results provided by equivalence testing now clearly indicate that the model is in fact not tenable, but one that only achieves fair fit. Based on an examination of these fit criteria, we would now conclude that the single-factor model is therefore not a plausible model for the analyzed data set. These results are of course quite expected, as the data-generating mechanism used two underlying latent factors and not one. Nevertheless, the presented example does demonstrate an instance where an incorrectly hypothesized one-factor model would not be rejected by conventional goodness-of-fit indexes but would be identified by equivalence testing as not a plausible model.

Next, fitting a two-factor model to the same data set (by having the first seven observed variables load on the first factor and the remaining five variables on the second factor), the following conventional goodness-of-fit indexes were obtained: $\chi^2(53)$ = 58.374, CFI = 0.996, and RMSEA = .013, with a 90% CI of [0, .030]. The complete model fit information provided by M*plus* is presented in Table 3. As can been seen by examining these results, model fit information can be considered indicative of a tenable model that achieves excellent fit (CFI > .99, p value > .05) or close fit (RMSEA < .05). These results are, of course, consistent with the data-generating mechanism that used two underlying latent factors, however, in realistic situations a researcher would not have the a priori benefit of the knowledge concerning the underlying structure of the data.

Entering again the appropriate values as input into the R program to determine the T-size indexes, the following output is provided:

```
#—————For T-size RMSEA————————————#;
Conventional RMSEA = 0.0130106
T-size RMSEA in equivalence testing = 0.02979863
#——————For T-size CFI——————————#;
Conventional CFI = 0.996345
T-size CFI in equivalence testing = 0.9749579
#————————————————————————————#;
```

TABLE 3
M*plus* Model Fit Information for Two-Factor Model

Number of free parameters	37
Log-likelihood	
H0 value	−10994.853
H1 value	−10965.666
Information criteria	
Akaike (AIC)	22063.705
Bayesian (BIC)	22226.392
Sample-size-adjusted BIC	22108.927
(n^* = (n + 2) /24)	
Chi-square test of model fit	
Value	58.374
Degrees of freedom	53
p value	0.2845
Root mean square error of approximation (RMSEA)	
Estimate	0.013
90% CI	[0.000, 0.030]
Probability RMSEA ≤ .05	1.000
CFI/TLI	
CFI	0.996
TLI	0.995
Chi-square test of model fit for the baseline model	
Value	1536.307
Degrees of freedom	66
p value	0.0000
Standardized root mean square residual (SRMR)	
Value	0.024

Note. CFI = comparative fit index; TLI = Tucker–Lewis Index.

Evaluating next these T-size results relative to excellent, close, fair, and mediocre or poor cutoff values as determined by the appropriate R program for equivalence testing, we obtain the following output:

```
#——————————————————————————————#;
CFI_e90      CFI_e92    CFI_e95    CFI_e99
[1,] 0.8666568  0.8905432  0.9271362  0.9777262
–poor– 0.867 –mediocre– 0.891 –fair– 0.927 –close– 0.978 –excellent–
RMSEA_e01 RMSEA_e05 RMSEA_e08 RMSEA_e10
[1,] 0.02841816  0.06045401  0.08943186  0.1090959
–excellent– 0.028 –close– 0.060 –fair– 0.089 –mediocre– 0.109 –poor–
#——————————————————————————————#;
```

The results provided by equivalence testing also indicate that the model is tenable and that it achieves close fit. Based on an examination of these fit criteria, as expected we would of course conclude that the two-factor model is a plausible model for the analyzed data set. It is important to mention here that, although equivalence testing might seem to provide an identical conclusion to that reached using conventional criteria, the implications of the results are very different. Specifically, for equivalence testing it is determined with 95% confidence that a misspecification in the model is no more than 0.02979863 (as measured by T-size RMSEA). Similarly, based on the T-size CFI, it is determined with 95% confidence that the population CFI is above 0.9749579. In contrast, with conventional criteria, no information is available about whether the model is correctly specified or about the degree of misspecification, if the model is in fact misspecified. Additionally, no claim with any desired level of probability can be made, concerning how far the CFI value is from the population value (Yuan et al., 2016).

CONCLUSION

Researchers are repeatedly confronted with the need to evaluate the goodness of fit of an SEM model believed to account for the complex multivariate relationships among observed and latent variables. Currently there appears to be a growing consensus in the literature that goodness of fit is most appropriately assessed using a variety of different perspectives (Yuan et al., 2016). Equivalence testing offers a novel yet rarely used approach for conducting model–data fit assessments in SEM. This didactic article illustrated the approach using two simple CFA examples in which a priori specified latent variable models with known structure were tested against data. The obtained results highlighted the tremendous benefits and additional insights gained by using equivalence testing when SEM is selected as the relevant methodology.

We advocate that whenever the goodness of fit of a model is being examined, in addition to considering conventional fit indexes, researchers should always examine the resulting values obtained by the equivalence testing approach. This is because, for the same data set and model tested, equivalence testing can provide more information than only using currently available conventional fit indexes. As indicated by Yuan et al. (2016), the obtained T-size values empower researchers to make a claim concerning the goodness of fit of an examined model with a degree of confidence giving SEM the desired property of being a scientific methodology.

REFERENCES

Arbuckle, J. L. (2012). *IBM SPSS AMOS 21 user's guide*. Chicago, IL: AMOS Development Corporation.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*, 815–824. doi:10.1016/j.paid.2006.09.018

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246. doi:10.1037/0033-2909.107.2.238

Bentler, P. M. (2006). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., … Fox, J. (2011). OpenMx: An open source extended structural equation modeling framework. *Psychometrika*, *76*, 306–317. doi:10.1007/s11336-010-9200-6

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230–258. doi:10.1177/0049124192021002005

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications and programming*. Mahwah, NJ: Erlbaum.

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. doi:10.1037/1082-989X.3.4.424

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 user's reference guide*. Chicago, IL: Scientific Software International.

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical processes. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling techniques* (pp. 315–353). Mahwah, NJ: Erlbaum.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320–341. doi:10.1207/s15328007sem1103_2

Muthén, L. K., & Muthén, B. O. (2014). *M*plus *user's guide*. Los Angeles, CA: Muthén & Muthén.

Pritikin, J. N., Hunter, M. D., & Boker, S. M. (2015). Modular open-source software for item factor analysis. *Educational and Psychological Measurement*, *75*, 458–474. doi:10.1177/0013164414554615

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. New York, NY: Routledge, Taylor & Francis.

Raykov, T., Marcoulides, G. A., & Tong, B. (2015). Do two or more multicomponent instruments measure the same construct? Testing construct congruence using latent variable modeling. *Educational and Psychological Measurement*. Advance online publication. doi: 10.1177/0013164415604705

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. doi:10.18637/jss.v048.i02

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, *28*, 1629–1646. doi:10.1177/014616702237645

Steiger, J. H., & Lind, J. M. (1980, July). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, *1*, 31–65. doi:10.1146/annurev.clinpsy.1.102803.144239

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Unpublished doctoral dissertation). University of California, Los Angeles, CA.

Yuan, K.-H., Chan, W., Marcoulides, G. A., & Bentler, P. M. (2016). Assessing structural equation models by equivalence testing with adjusted fit indexes. *Structural Equation Modeling*, *23*, 319–330. doi:10.1080/10705511.2015.1065414