ESEM, CFA, and Somewhere In-Between: The Effect of Measurement Quality on Model Fit Sensitivity

Tim Konold, Ph.D., University of Virginia

Liz Sanders, Ph.D. University of Washington

University of Virginia | School of Education and Human Development

University of Washington

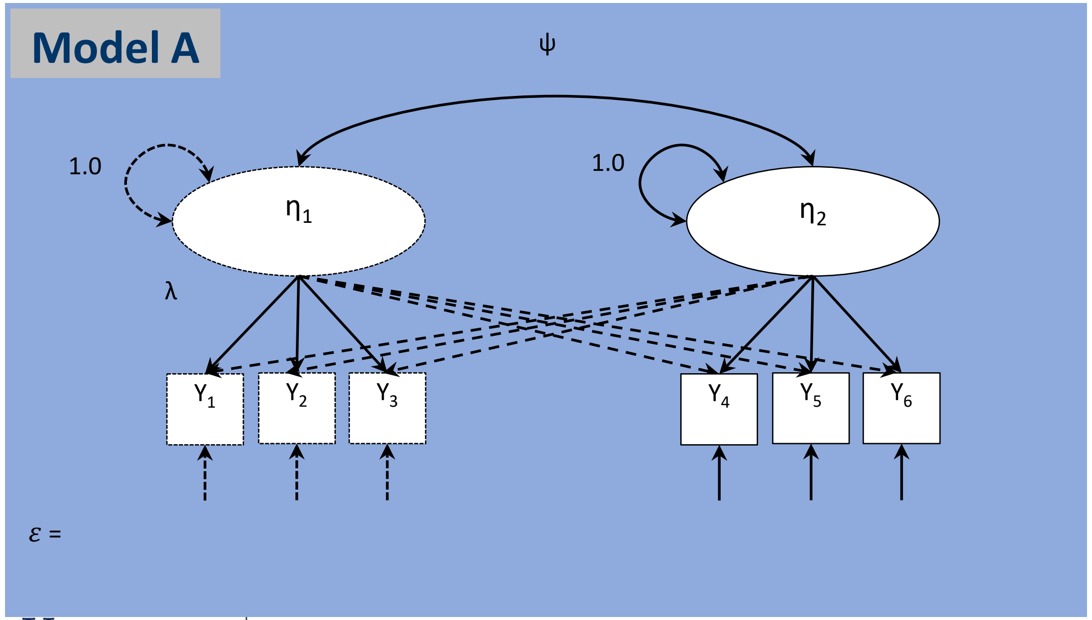
ESEM, CFA, and Somewhere In-Between: The Effect of ~~Measurement Quality~~ Cross Loading Sign on Model Fit Sensitivity

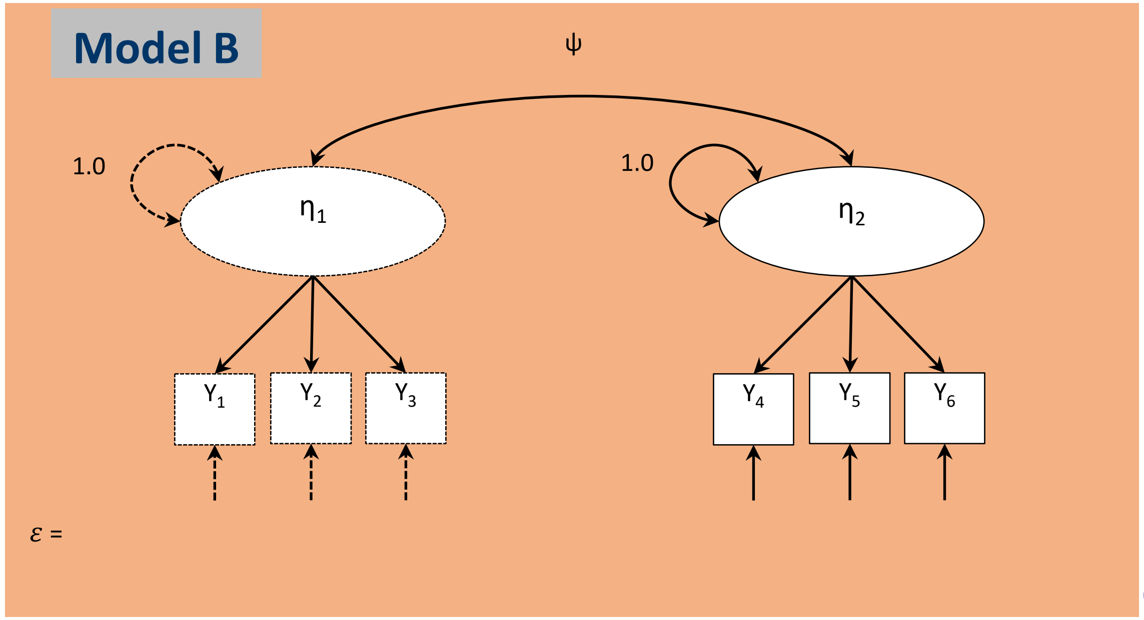
Tim Konold, Ph.D., University of Virginia
Liz Sanders, Ph.D. University of Washington

University of Virginia | School of Education and Human Development

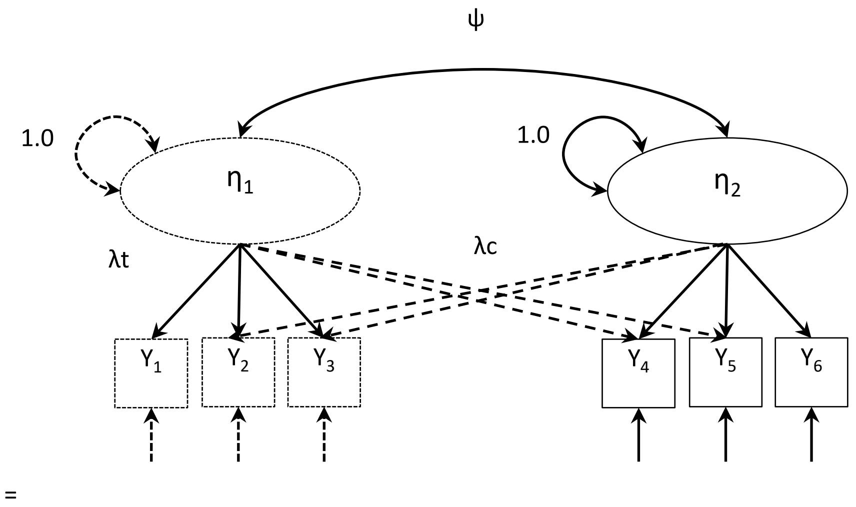
University of Washington

Introduction

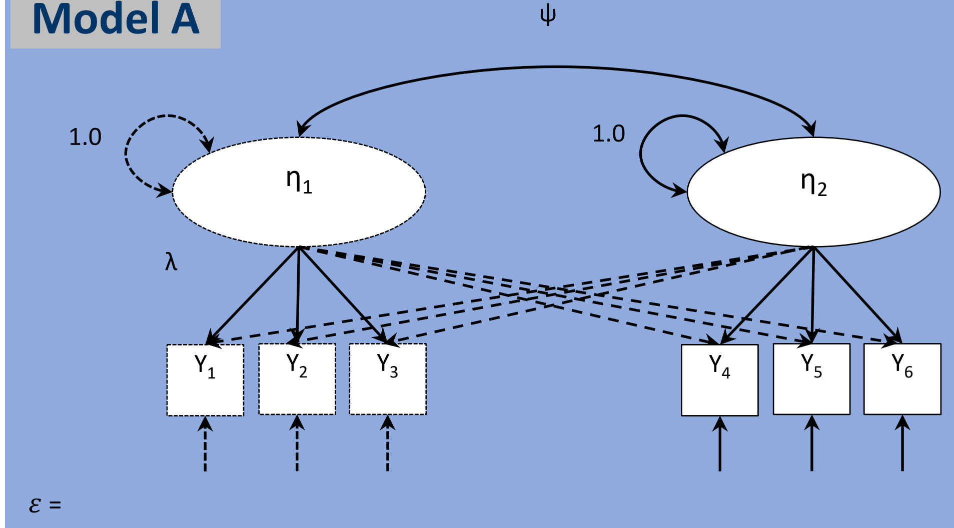




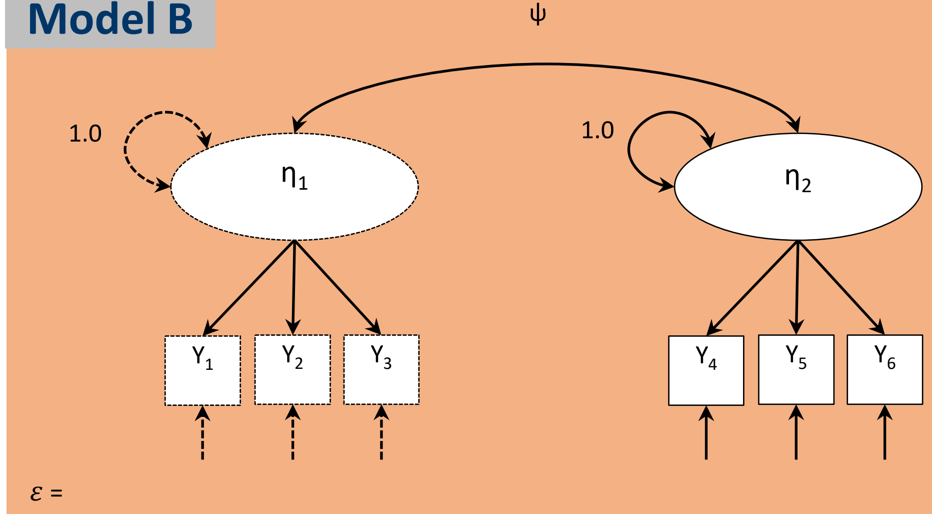
Introduction



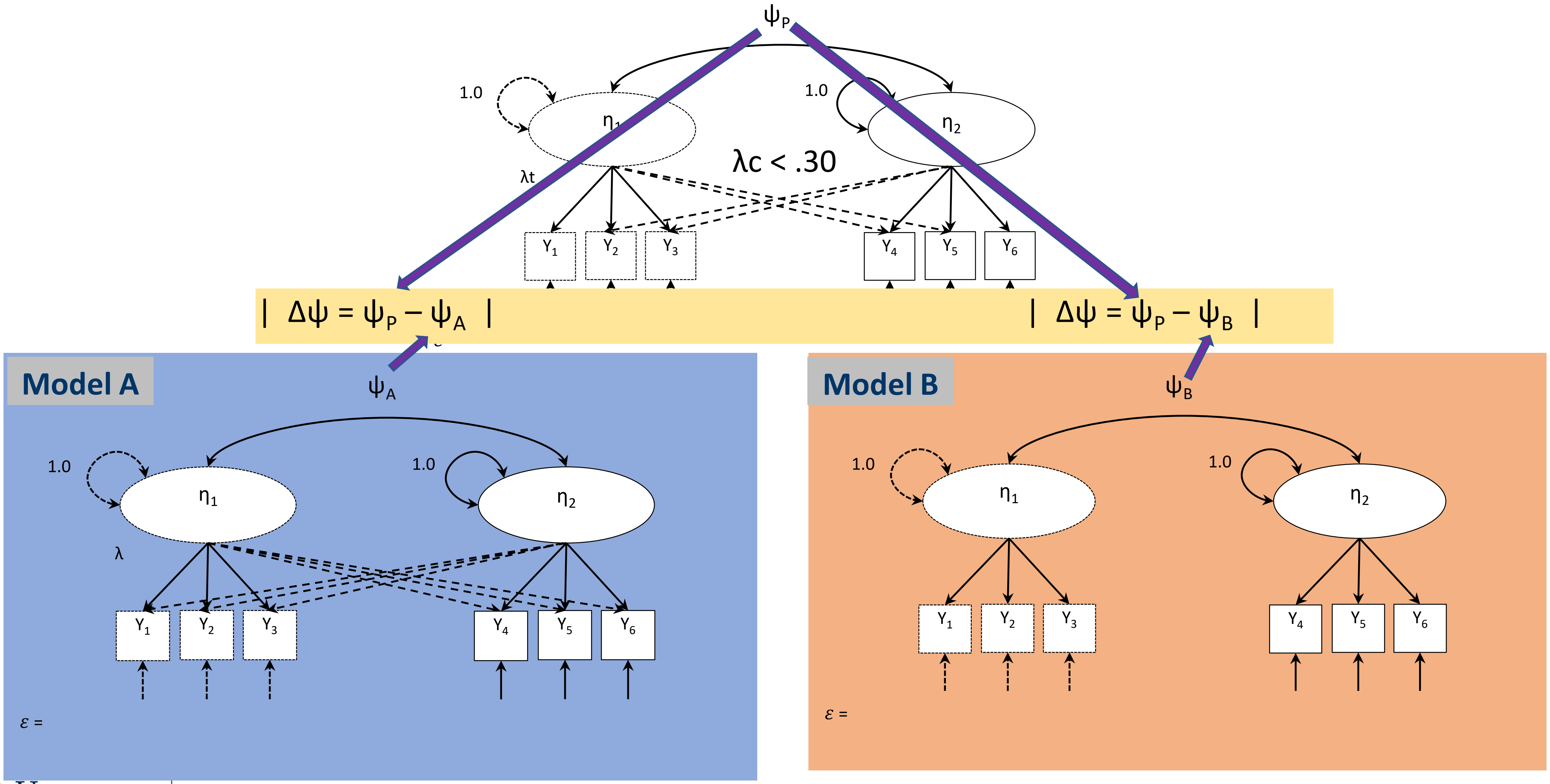
Model A



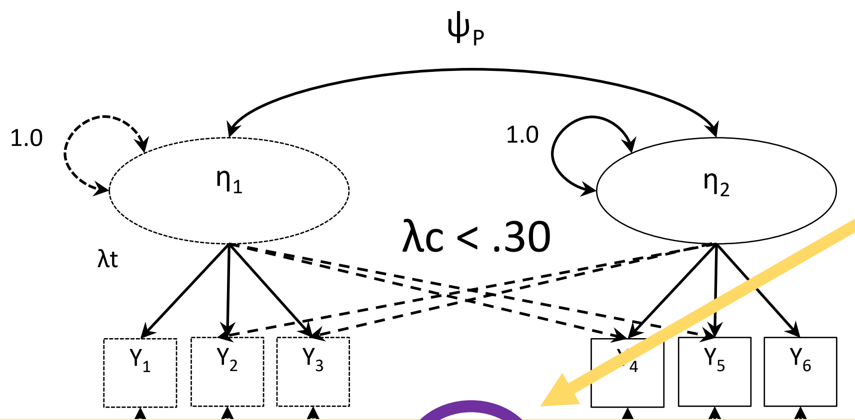
Model B



Introduction



Introduction



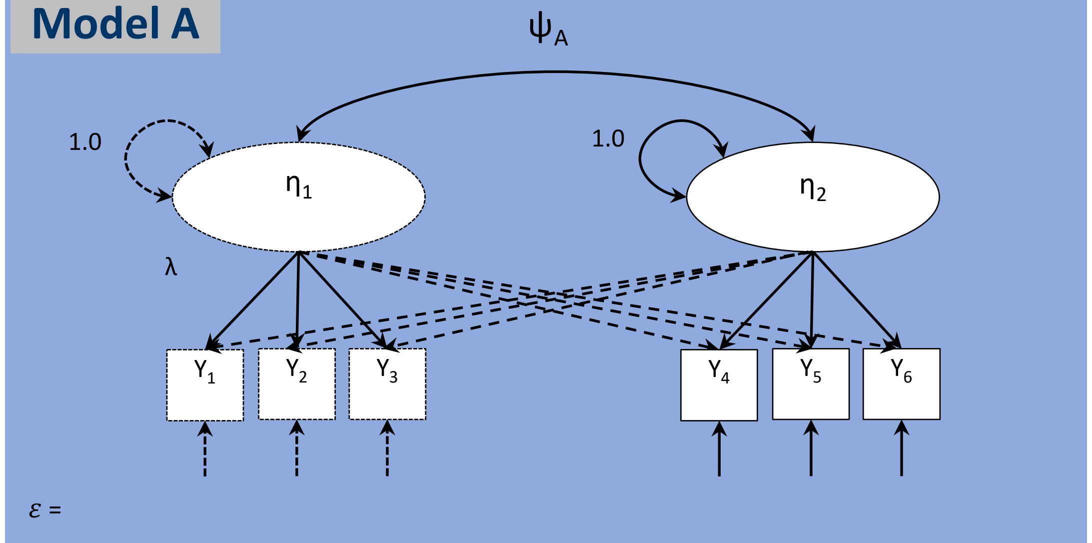
Asparouhov & Muthen (2009)

Marsh et al., (2009…2020)

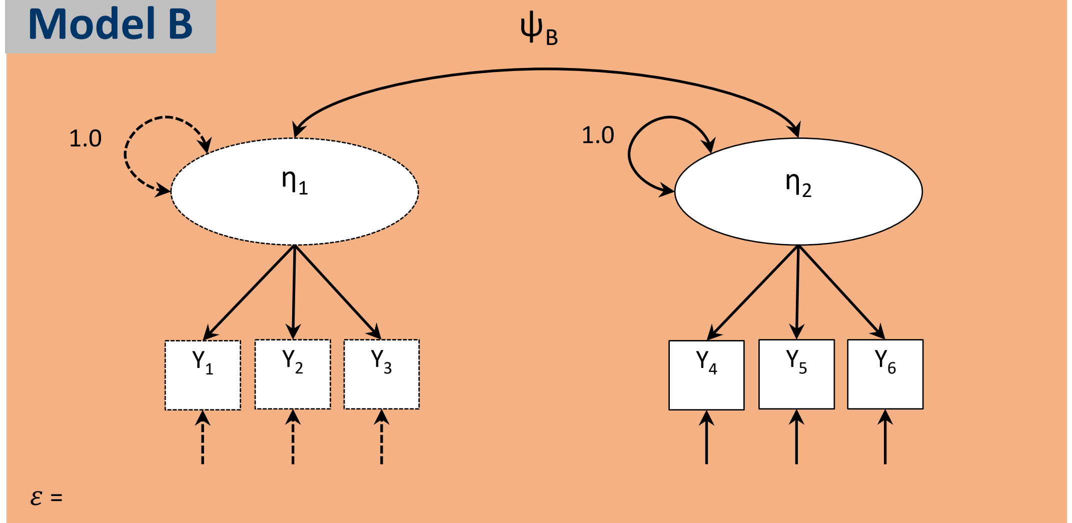
Steenkamp & Maydeu-Olivares, (2023)

$| \Delta\psi = \psi_P - \psi_A |$ $<$ $| \Delta\psi = \psi_P - \psi_B |$

Model A



Model B



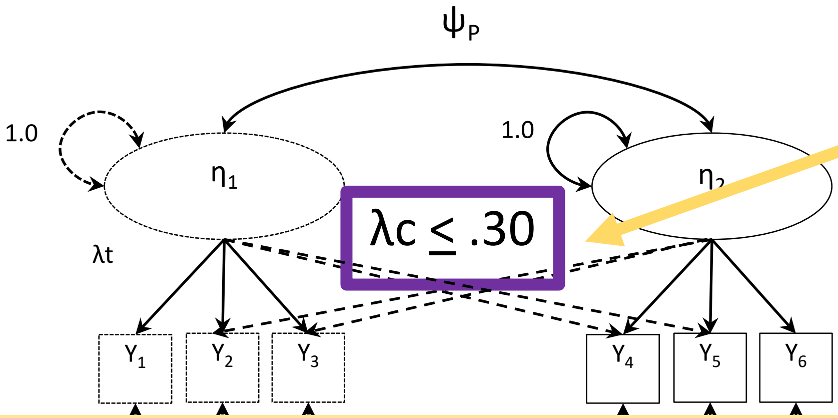
Introduction

Tabachnick & Fidell, 2007

Cudeck & O'Dell, 1994

van Prooijen & van der Kloot, 2001

Grice, 2010

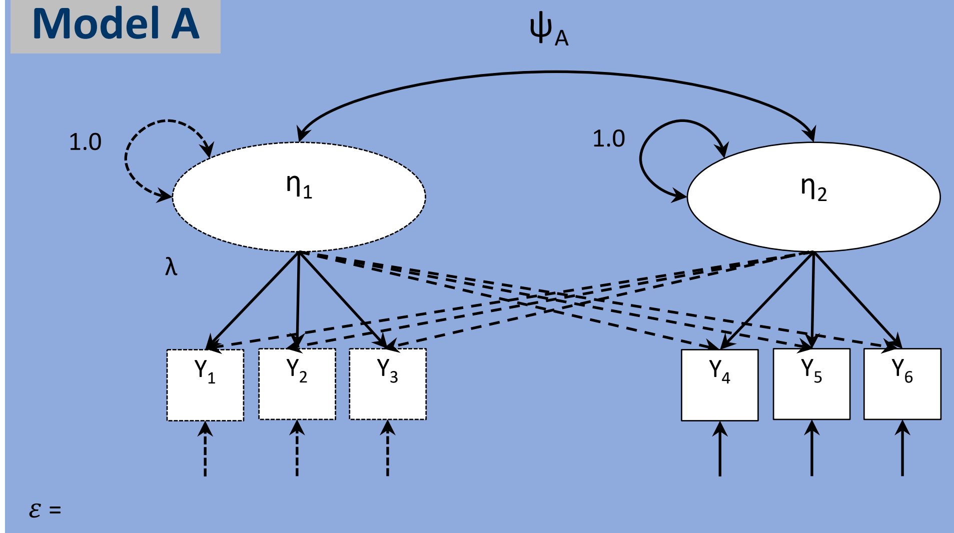


YAY!

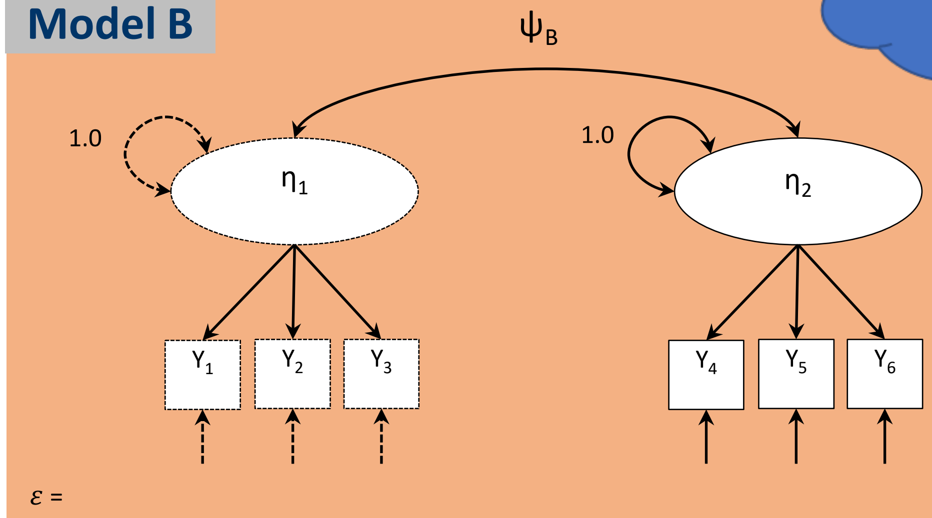
$$| \Delta\psi = \psi_P - \psi_A | \quad < \quad | \Delta\psi = \psi_P - \psi_B |$$

That's not fair!

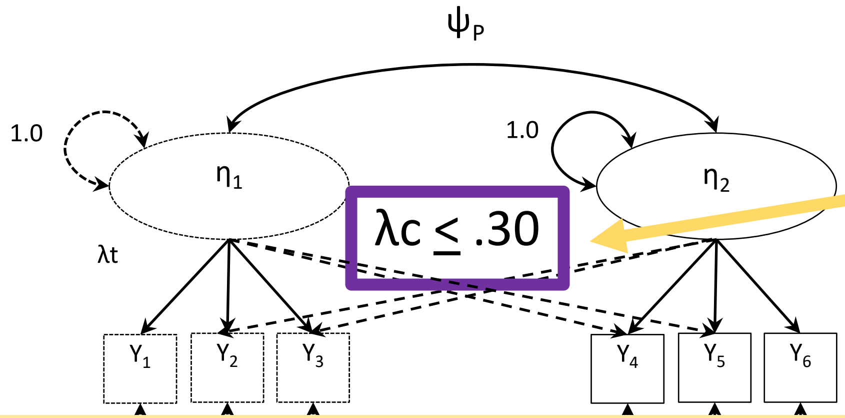
Model A



Model B



Introduction



Hsu, et al. (2014)

λc of $|.13|$ → 0; → biased ψ_B

Steenkamp & Maydeu-Olivares (2023)

λc of $|.2|$ → 0;

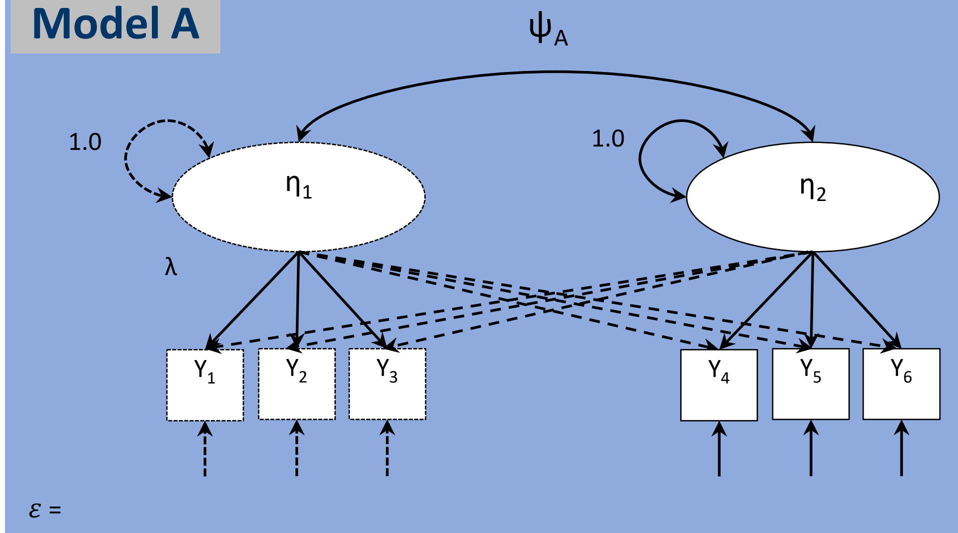
→ Non-negligible impact on ψ_B

λc of $|.3|$ → 0;

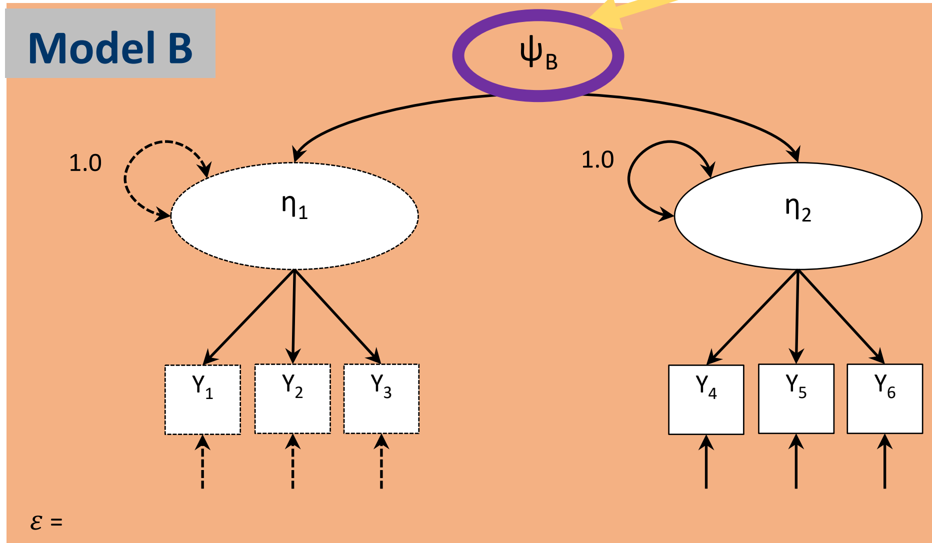
→ "appreciable" bias in ψ_B

$|\ \Delta\psi = \psi_P - \psi_A\ |$ < | Δ

Model A



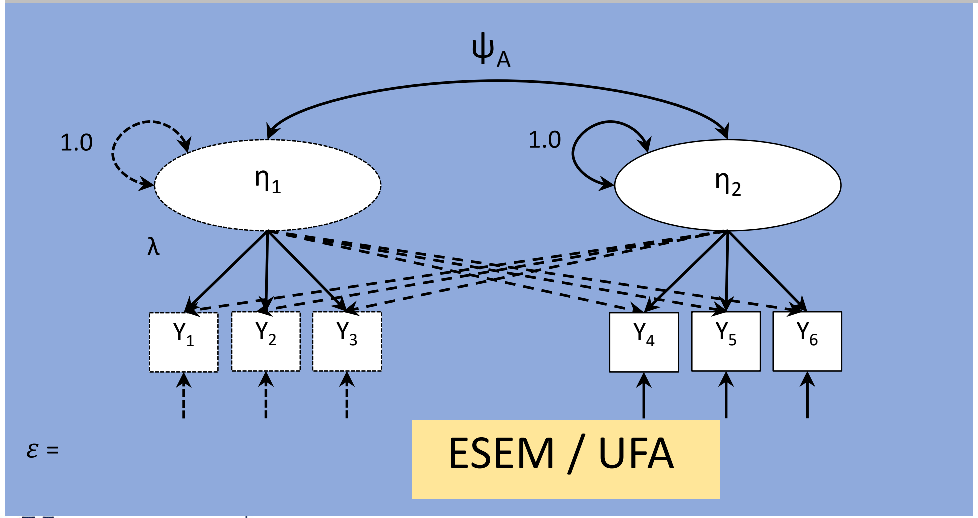
Model B

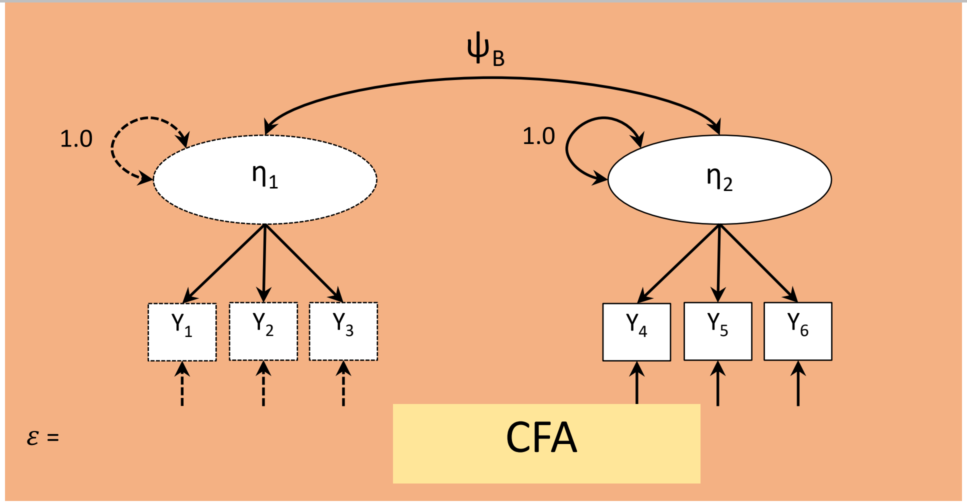


Introduction

What to do? How do we balance these advantages?

ESEM / UFA	**CFA**
Better "Fit", less likely to reject 'trivially' mis-specified models	**Parsimony**
More realistic evaluation of simple structure	
Less biased structural relationships	
Allows for a more approximate hypothesis of the measurement structure	
~~Single solution that is invariant to type of rotation~~	

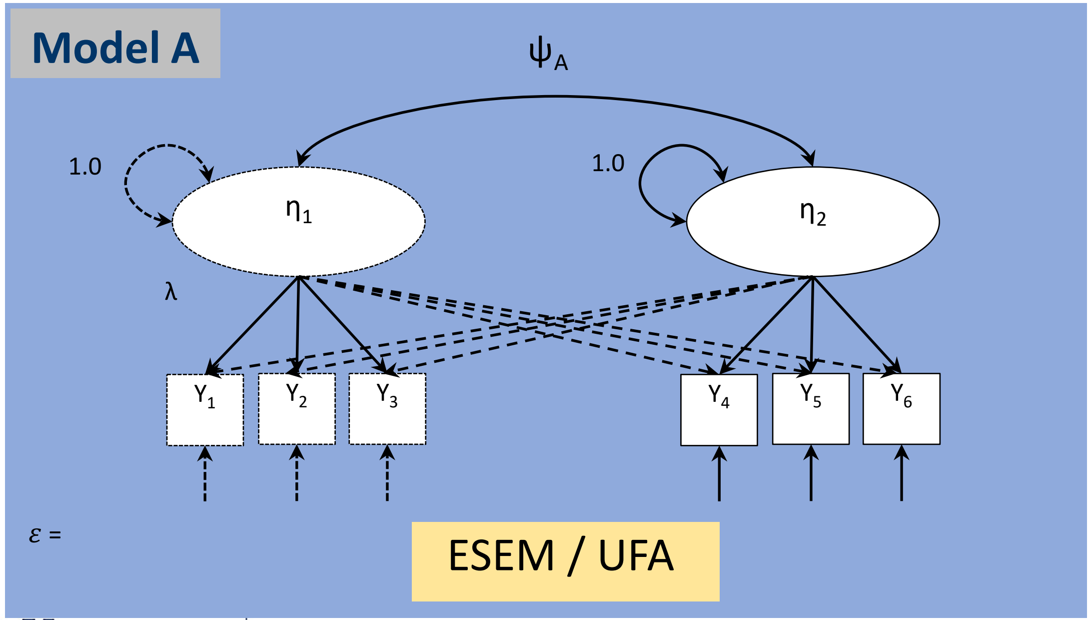


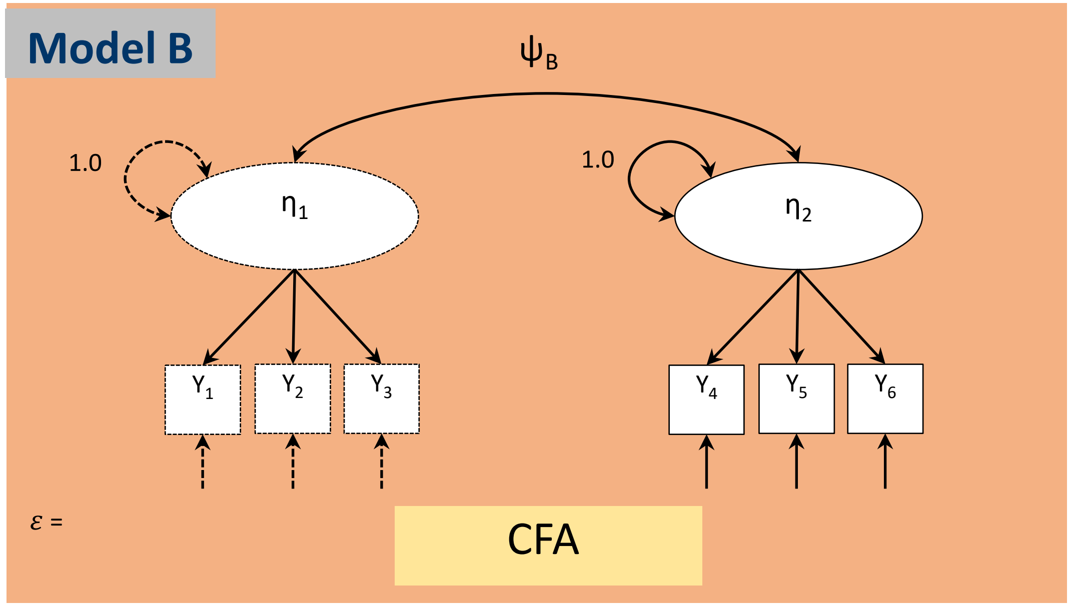


Introduction

A strategy (Steenkamp & Maydeu-Olivares, 2023)

~Estimate both UFA and CFA models
~Compare model estimates (e.g., factor correlations, factor loadings)
~IFF LRT rejects the CFA model…
~Evaluate overall fit of each (LRT, CFI, RMSEA)
~Evaluate nested model comparisons (ΔLRT, Δ CFI, ΔRMSEA)

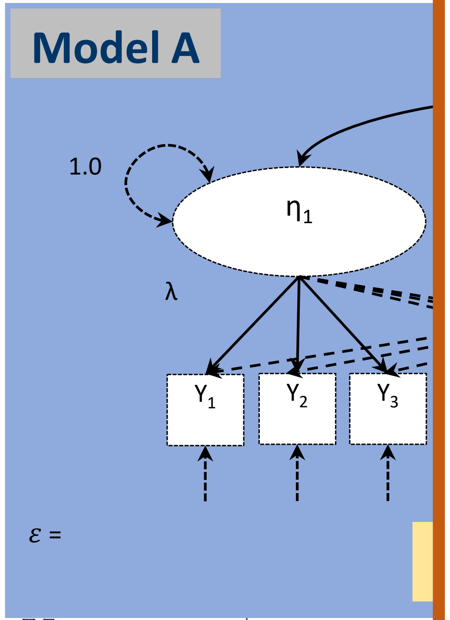




Introduction

A strategy (Steenkamp & Maydeu-Olivares, 2023)

~Estimate both UFA and CFA models
~Compare model estimates (e.g., factor correlations, factor loadings)
~Evaluate overall fit of each (LRT, CFI, RMSEA)
~Evaluate nested model comparisons (ΔLRT, ΔCFI, ΔRMSEA)



Remain the most frequently reported measures of fit in SEM (Jackson et al., 2009)

Have been emphasized in recent methodological evaluations of UFA (Steenkamp & Maydeu-Olivares, 2023; Marsh, et al., 2020)

with van Zyl and Klooster (2022) noting that in evaluations of UFA "the CFI, TLI, and RMSEA should always be reported and used as the primary criterion for both establishing model fit and to discriminate between models" (p. 9)

The Current Study

- Briefly revisit structural bias when non-zero cross loadings in the population are constrained to zero
- Examine the degree to which GFI and ΔGFI measures are sensitive to what has been considered ignorable cross loadings
- Examine GFI and ΔGFI (in comparison to ESEM/UFA specification) measures relative to the CFA models that give rise to structural relationship bias and those that do not

Evaluating Fit

Stand-Alone Models

$$LR = (N - 1)Fml \sim \chi^2_{df\,H}$$

$$CFI = 1 - \frac{\max[0, (\chi^2_H - df_H)]}{\max[0, (\chi^2_B - df_B)]}$$

$$RMSEA = \sqrt{\max\left[0, \frac{\chi^2_H - df_H}{df_H \times (N)}\right]}$$

Nested Models

$$\Delta LR = LR_B - LR_A \sim \chi^2_{dfB\,-\,dfA}$$

$$\Delta CFI = CFI_A - CFI_B$$

$$\Delta RMSEA = RMSEA_A - RMSEA_B$$

$$RMSEA_D = \sqrt{\frac{\Delta LR - df_D}{df_D(N-1)}}$$

UNIVERSITY of VIRGINIA | SCHOOL of EDUCATION and HUMAN DEVELOPMENT

UNIVERSITY of WASHINGTON

$RMSEA_D$ (Browne & Du, 1992), recently re-introduced (Savalei et al., 2023)

~Integration of ΔLR and ΔRMSEA (replaces NCP in RMSEA with ΔLR-df_D; and model df with df_D)
~Purports to over come limitations of ΔLR (too restrictive) and ΔRMSEA (too forgiving)
~Said to better focus on changes in the two nested models (relative to df) rather than common aspects
~As n→N, $RMSEA_D$ → population value; rather than being overpowered for trivial differences
~On same scale as RMSEA, and provides for CIs

$$LR = (N-1)Fml \sim \chi^2_{df\,H}$$

$$\Delta LR = LR_B - LR_A \sim \chi^2_{dfB\,-\,dfA}$$

$$CFI = 1 - \frac{\max[0,(\chi^2_H - df_H)]}{\max[0,(\chi^2_B - df_B)]}$$

$$\Delta CFI = CFI_A - CFI_B$$

$$RMSEA = \sqrt{\max\left[0, \frac{\chi^2_H - df_H}{df_H \times (N)}\right]}$$

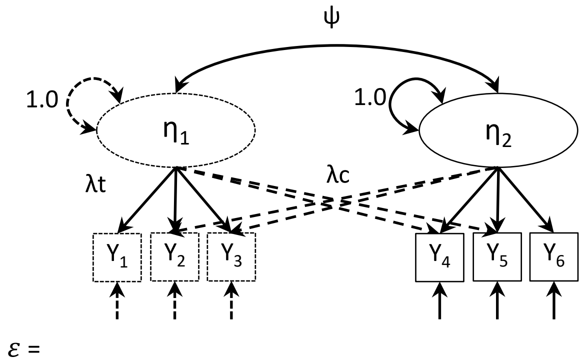
$$\Delta RMSEA = RMSEA_A - RMSEA_B$$

$$RMSEA_D = \sqrt{\frac{\Delta LR - df_D}{df_D(N-1)}}$$

UNIVERSITY of VIRGINIA | SCHOOL of EDUCATION and HUMAN DEVELOPMENT

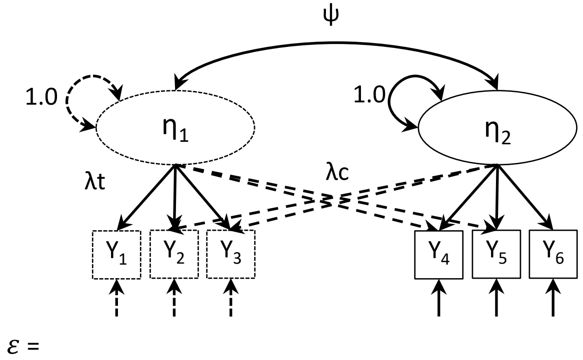
UNIVERSITY of WASHINGTON

Population Model

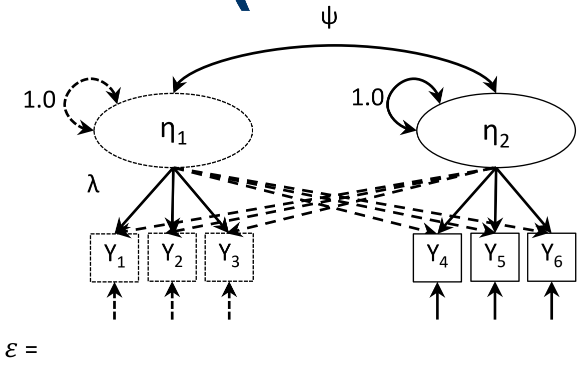


Two-Factor Models with unit normal factors and indicator scales

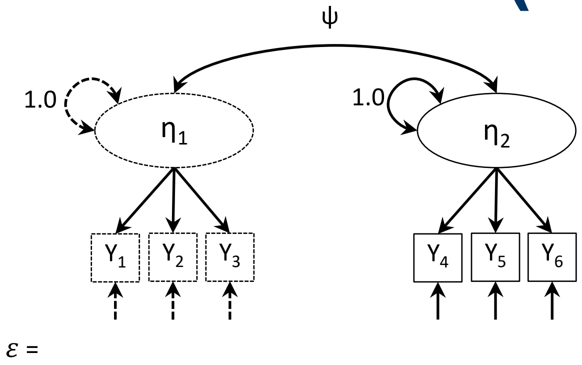
Correct (CS)



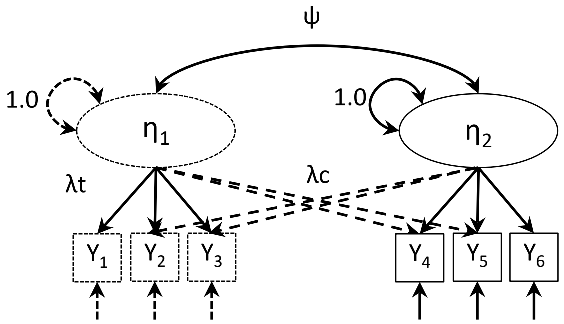
UFA (aka ESEM)



ICM-CFA (IS)



Population Model



With an Emphasis on Measurement Quality of the Target Indicators

6-, 12- and 24-indicator models; *crossed with*

Target Loadings: λt = .40, .50, .60, 70, .80; *crossed with*

Cross Loadings: λc = .00, .15, .30; *nearly crossed with*

Cross Loading Saturation: 33%, 67% *crossed with*

Cross Loading Sign: Positive, Negative, Mixed[1] *nearly crossed with*

Factor Correlations: $\varphi_{1,2}$ *= 0, .1, .2, .3, .4, and .5;* *crossed with*

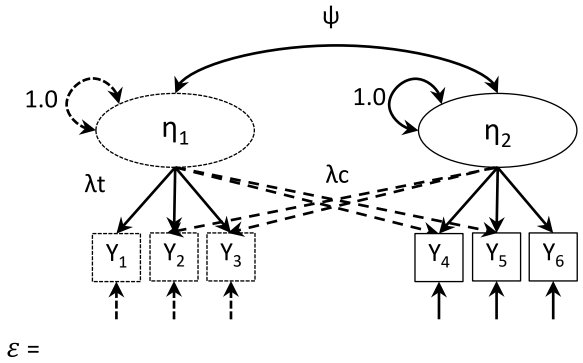
Sample Sizes: N = 200, 500, 1000

(*N* = 1,000 data sets were generated for the 4,860 conditions)

1. Absolute cross-loading magnitudes were held constant across signs (e.g., ± .15 was only paired with ± .15).

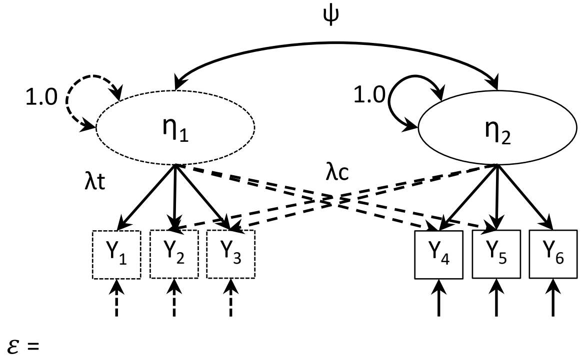
Simulation Conditions

Population Model

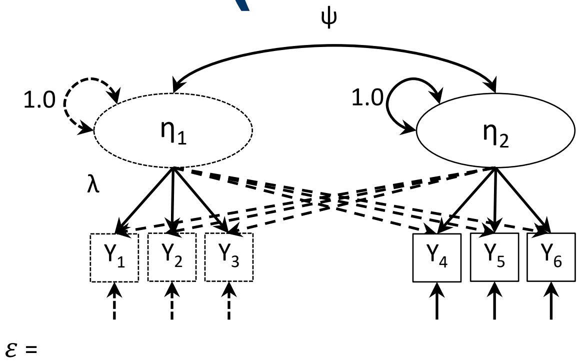


- Data were generated and analyzed using *Mplus 8* within the *MplusAutomation R* package
- All models were estimated with ML
- Results today based on UFA with Target rotation

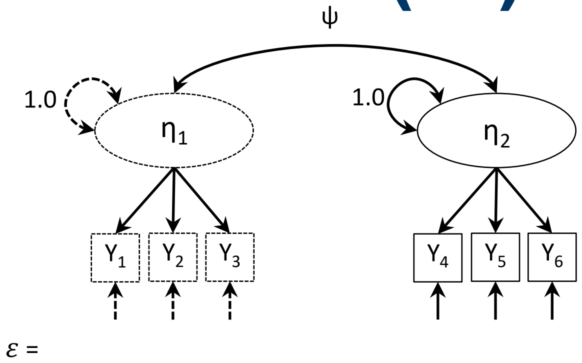
Correct (CS)



UFA (aka ESEM)



CFA (IS)



Results: Factor Correlation Bias

$= \psi_{Population} - \psi_{estimated\ model}$

Factor Correlation Raw Bias across Models for Selected Conditions: N = 1,000 Sample Size

						33% of Indicators Cross-Load									67% of Indicators Cross-Load								
						All Positive CLs			All Negative CLs			Mix Pos/Neg CLs			All Positive CLs			All Negative CLs			Mix Pos/Neg CLs		
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	Pop Cor	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA
1000	6	.0	.4	.15	.0	.00	.22	.27	.00	-.22	-.27	.00	.00	.00	.02	.39	.52	-.02	-.38	-.52	.00	.00	-.01
				.30	.0	.01	.38	.53	-.01	-.38	-.52	.00	.00	.00	.05	.44	.89	-.04	-.44	-.89	.00	.00	-.02
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.27	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.29	.00	-.24	-.29	.00	.00	.00	.00	.43	.56	.00	-.43	-.56	.00	.00	.00
		.5	.4	.15	.5	.00	.03	.17	.00	-.27	-.19	.00	-.09	.06	-.02	.07	.31	-.01	-.45	-.43	.01	-.11	.12
				.30	.5	-.01	.07	.29	.00	-.50	-.65	.00	-.13	.14	-.06	.03	.46	-.06	-.69	-1.19	.01	-.19	.33
			.8	.15	.5	.00	.08	.10	.00	-.11	-.08	.00	-.01	.06	.00	.16	.19	.00	-.21	-.17	.00	-.02	.11
				.30	.5	.00	.14	.21	.00	-.23	-.14	.00	-.03	.16	.00	.26	.37	.00	-.45	-.40	.00	-.06	.27
	24	.0	.4	.15	.0	.00	.23	.26	.00	-.23	-.26	.00	.00	.00	.00	.42	.50	.00	-.42	-.50	.00	.00	.00
				.30	.0	.00	.40	.51	.00	-.40	-.51	.00	.00	.00	.00	.59	.89	.00	-.59	-.88	.00	.00	.00
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.26	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.28	.00	-.24	-.28	.00	.00	.00	.00	.43	.54	.00	-.43	-.54	.00	.00	.00
		.5	.4	.15	.5	.00	.12	.17	.00	-.24	-.20	.00	-.05	.07	.00	.23	.31	.00	-.45	-.43	.00	-.07	.13
				.30	.5	.00	.19	.29	.00	-.50	-.51	.00	-.10	.23	.01	.28	.46	.00	-.76	-1.11	.00	-.16	.40
			.8	.15	.5	.00	.08	.10	.00	-.11	-.09	.00	-.01	.04	.00	.16	.19	.00	-.21	-.19	.00	-.02	.09
				.30	.5	.00	.14	.21	.00	-.23	.32	.00	-.03	.19	.00	.26	.37	.00	-.45	-.38	.00	-.06	.35

Results: Factor Correlation Bias $= \psi_{Population} - \psi_{estimated\ model}$

Factor Correlation Raw Bias across Models for Selected Conditions: N = 1,000 Sample Size

						33% of Indicators Cross-Load									67% of Indicators Cross-Load								
						All Positive CLs			All Negative CLs			Mix Pos/Neg CLs			All Positive CLs			All Negative CLs			Mix Pos/Neg CLs		
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	Pop Cor	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA	CFA (CS)	UFA	CFA
1000	6	.0	.4	.15	.0	.00	.22	.27	.00	-.22	-.27	.00	.00	.00	.02	.39	.52	-.02	-.38	-.52	.00	.00	-.01
				.30	.0	.01	.38	.53	-.01	-.38	-.52	.00	.00	.00	.05	.44	.89	-.04	-.44	-.89	.00	.00	-.02
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.27	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.29	.00	-.24	-.29	.00	.00	.00	.00	.43	.56	.00	-.43	-.56	.00	.00	.00
		.5	.4	.15	.5	.00	.03	.17	.00	-.27	-.19	.00	-.09	.06	-.02	.07	.31	-.01	-.45	-.43	.01	-.11	.12
				.30	.5	-.01	.07	.29	.00	-.50	-.65	.00	-.13	.14	-.06	.03	.46	-.06	-.69	-1.19	.01	-.19	.33
			.8	.15	.5	.00	.08	.10	.00	-.11	-.08	.00	-.01	.06	.00	.16	.19	.00	-.21	-.17	.00	-.02	.11
				.30	.5	.00	.14	.21	.00	-.23	-.14	.00	-.03	.16	.00	.26	.37	.00	-.45	-.40	.00	-.06	.27
	24	.0	.4	.15	.0	.00	.23	.26	.00	-.23	-.26	.00	.00	.00	.00	.42	.50	.00	-.42	-.50	.00	.00	.00
				.30	.0	.00	.40	.51	.00	-.40	-.51	.00	.00	.00	.00	.59	.89	.00	-.59	-.88	.00	.00	.00
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.26	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.28	.00	-.24	-.28	.00	.00	.00	.00	.43	.54	.00	-.43	-.54	.00	.00	.00
		.5	.4	.15	.5	.00	.12	.17	.00	-.24	-.20	.00	-.05	.07	.00	.23	.31	.00	-.45	-.43	.00	-.07	.13
				.30	.5	.00	.19	.29	.00	-.50	-.51	.00	-.10	.23	.01	.28	.46	.00	-.76	-1.11	.00	-.16	.40
			.8	.15	.5	.00	.08	.10	.00	-.11	-.09	.00	-.01	.04	.00	.16	.19	.00	-.21	-.19	.00	-.02	.09
				.30	.5	.00	.14	.21	.00	-.23	.32	.00	-.03	.19	.00	.26	.37	.00	-.45	-.38	.00	-.06	.35

Results

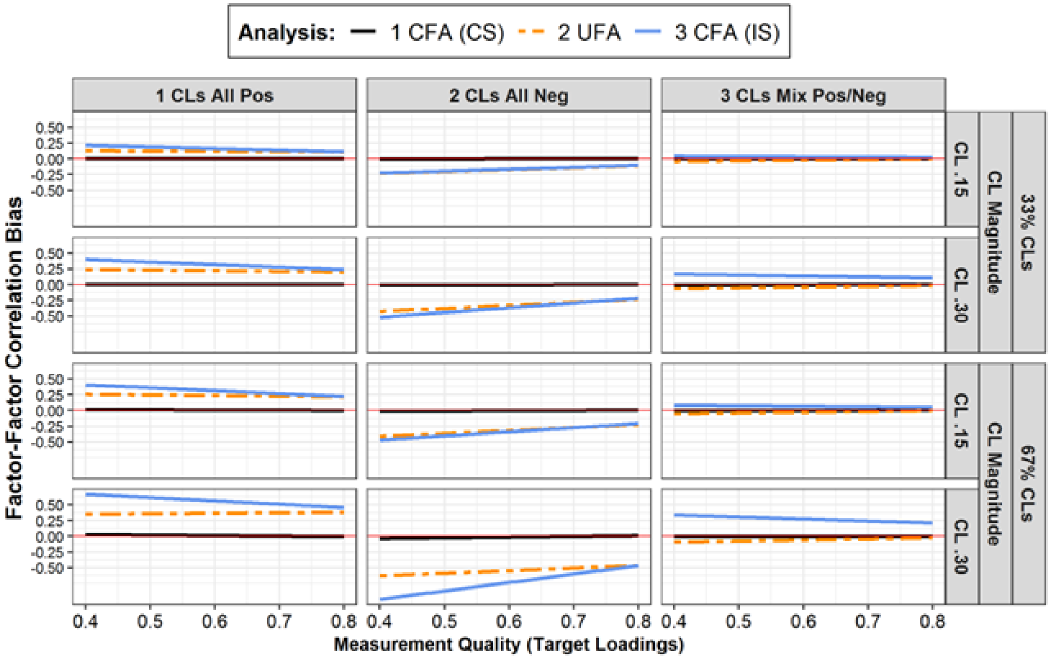
Bias = $\psi_{Population} - \psi_{estimated\ model}$

Factor Correlation Raw Bias across Models for Selected Conditions: N = 1,000 Sample Size

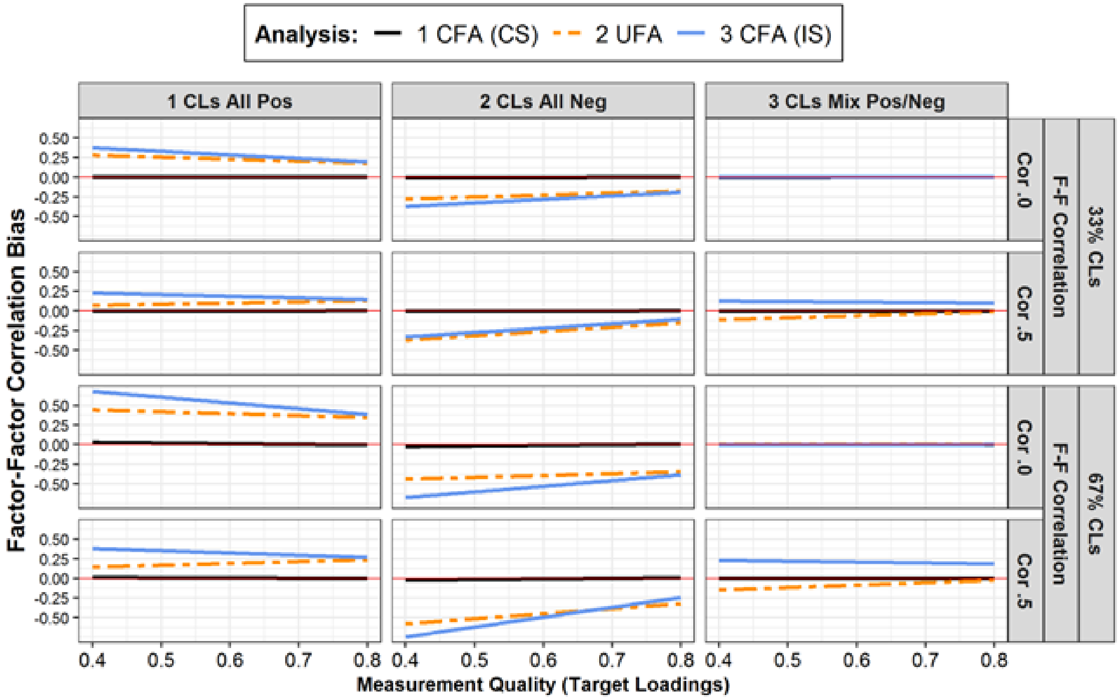
						33% of Indicators Cross-Load									67% of Indicators Cross-Load								
						All Positive CLs			All Negative CLs			Mix Pos/Neg CLs			All Positive CLs			All Negative CLs			Mix Pos/Neg CLs		
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	Pop Cor	CFA (CS)	UFA	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (CS)	UFA	CFA (IS)
1000	6	.0	.4	.15	.0	.00	.22	.27	.00	-.22	-.27	.00	.00	.00	.02	.39	.52	-.02	-.38	-.52	.00	.00	.00
				.30	.0	.01	.38	.53	-.01	-.38	-.52	.00	.00	.00	.05	.44	.89	-.04	-.44	-.89	.00	.00	.01
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.27	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.29	.00	-.24	-.29	.00	.00	.00	.00	.43	.56	.00	-.43	-.56	.00	.00	.00
		.5	.4	.15	.5	.00	.03	.67	.00	-.27	.31	.00	-.09	.56	-.02	.07	.81	-.01	-.45	.07	.01	-.11	.62
				.30	.5	-.01	.07	.79	.00	-.50	-.15	.00	-.13	.64	-.06	.03	.96	-.06	-.69	-.69	.01	-.19	.83
			.8	.15	.5	.00	.08	.60	.00	-.11	.42	.00	-.01	.56	.00	.16	.69	.00	-.21	.33	.00	-.02	.61
				.30	.5	.00	.14	.71	.00	-.23	.36	.00	-.03	.66	.00	.26	.87	.00	-.45	.10	.00	-.06	.77
	24	.0	.4	.15	.0	.00	.23	.26	.00	-.23	-.26	.00	.00	.00	.00	.42	.50	.00	-.42	-.50	.00	.00	.00
				.30	.0	.00	.40	.51	.00	-.40	-.51	.00	.00	.00	.00	.59	.89	.00	-.59	-.88	.00	.00	.00
			.8	.15	.0	.00	.12	.13	.00	-.12	-.13	.00	.00	.00	.00	.24	.26	.00	-.24	-.26	.00	.00	.00
				.30	.0	.00	.24	.28	.00	-.24	-.28	.00	.00	.00	.00	.43	.54	.00	-.43	-.54	.00	.00	.00
		.5	.4	.15	.5	.00	.12	.67	.00	-.24	.30	.00	-.05	.57	.00	.23	.81	.00	-.45	.07	.00	-.07	.63
				.30	.5	.00	.19	.79	.00	-.50	-.01	.00	-.10	.73	.01	.28	.96	.00	-.76	-.61	.00	-.16	.90
			.8	.15	.5	.00	.08	.60	.00	-.11	.41	.00	-.01	.54	.00	.16	.69	.00	-.21	.31	.00	-.02	.59
				.30	.5	.00	.14	.71	.00	-.23	.32	.00	-.03	.69	.00	.26	.87	.00	-.45	.12	.00	-.06	.85

Factor-Factor *Correlation Bias across Analysis*

CL Magnitude



F-F Correlation



$$= \psi_{Population} - \psi_{CFA}$$

Partial Eta-Squared Values for Raw CFA Factor-Factor Correlation Bias

Population Design Conditions	Raw FF R Bias
Main Effects	
N	0.00
Model Size	0.01
Factor Correlation (FC)	0.21
Target Loadings (TL)	0.20
Cross Loading (CL) Value	0.20
CL Sign	1.00
CL Saturation	0.01
Two-Way Interactions	
N * Model Size	0.00
N * FC	0.00
N * TL	0.00
N * CL Value	0.00
N * CL Sign	0.00
N * CL Saturation	0.00
Model Size * FC	0.01
Model Size * TL	0.02
Model Size * CL Value	0.03
Model Size * CL Sign	0.04
Model Size * CL Saturation	0.00
FC * TL	0.26
FC * CL Value	0.12
FC * CL Sign	0.85
FC * CL Saturation	0.07
TL * CL Value	0.11
TL * CL Sign	0.92
TL * CL Saturation	0.04
CL Value * CL Sign	0.96
CL Value * CL Saturation	0.00
CL Sign * CL Saturation	0.95

RMSEA

RMSEA Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

					All Positive Cross-Loadings				All Negative Cross-Loadings				Mix of Pos/Neg Cross-Loadings			
					Analysis			UFA vs.	Analysis			UFA vs.	Analysis			UFA vs.
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)
1000	6	.0	.4	.15	.008	.008	.026	-.018	.008	.008	.025	-.017	.009	.009	.066	-.057
				.30	.007	.007	.029	-.022	.006	.006	.028	-.022	.010	.010	.122	-.113
			.8	.15	.010	.010	.090	-.080	.009	.009	.089	-.080	.010	.010	.162	-.153
				.30	.010	.010	.164	-.154	.009	.009	.163	-.154	.010	.010	.326	-.316
		.5	.4	.15	.006	.006	.014	-.008	.008	.008	.030	-.022	.009	.009	.037	-.028
				.30	.004	.004	.015	-.011	.007	.007	.043	-.036	.009	.009	.065	-.056
			.8	.15	.010	.010	.071	-.061	.009	.009	.076	-.067	.010	.010	.123	-.113
				.30	.010	.010	.158	-.148	.009	.009	.155	-.146	.010	.010	.211	-.201
	24	.0	.4	.15	.004	.004	.018	-.014	.004	.004	.018	-.014	.004	.004	.035	-.031
				.30	.004	.004	.025	-.021	.004	.004	.025	-.021	.004	.004	.072	-.068
			.8	.15	.004	.004	.035	-.031	.004	.004	.035	-.031	.004	.004	.063	-.059
				.30	.004	.004	.069	-.065	.004	.004	.069	-.065	.004	.004	.131	-.127
		.5	.4	.15	.004	.004	.011	-.007	.004	.004	.018	-.014	.004	.004	.026	-.022
				.30	.004	.004	.014	-.010	.004	.004	.033	-.029	.004	.004	.044	-.040
			.8	.15	.004	.004	.030	-.026	.004	.004	.031	-.027	.004	.004	.053	-.049
				.30	.004	.004	.059	-.055	.004	.004	.062	-.058	.004	.004	.110	-.106

RMSEA

RMSEA Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	All Positive Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	All Negative Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	Mix of Pos/Neg Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)
1000	6	.0	.4	.15	.008	.008	.026	-.018	.008	.008	.025	-.017	.009	.009	.066	-.057
				.30	.007	.007	.029	-.022	.006	.006	.028	-.022	.010	.010	.122	-.113
			.8	.15	.010	.010	.090	-.080	.009	.009	.089	-.080	.010	.010	.162	-.153
				.30	.010	.010	.164	-.154	.009	.009	.163	-.154	.010	.010	.326	-.316
		.5	.4	.15	.006	.006	.014	-.008	.008	.008	.030	-.022	.009	.009	.037	-.028
				.30	.004	.004	.015	-.011	.007	.007	.043	-.036	.009	.009	.065	-.056
			.8	.15	.010	.010	.071	-.061	.009	.009	.076	-.067	.010	.010	.123	-.113
				.30	.010	.010	.158	-.148	.009	.009	.155	-.146	.010	.010	.211	-.201
	24	.0	.4	.15	.004	.004	.018	-.014	.004	.004	.018	-.014	.004	.004	.035	-.031
				.30	.004	.004	.025	-.021	.004	.004	.025	-.021	.004	.004	.072	-.068
			.8	.15	.004	.004	.035	-.031	.004	.004	.035	-.031	.004	.004	.063	-.059
				.30	.004	.004	.069	-.065	.004	.004	.069	-.065	.004	.004	.131	-.127
		.5	.4	.15	.004	.004	.011	-.007	.004	.004	.018	-.014	.004	.004	.026	-.022
				.30	.004	.004	.014	-.010	.004	.004	.033	-.029	.004	.004	.044	-.040
			.8	.15	.004	.004	.030	-.026	.004	.004	.031	-.027	.004	.004	.053	-.049
				.30	.004	.004	.059	-.055	.004	.004	.062	-.058	.004	.004	.110	-.106

Across the 4,320 design conditions, average (across 1,000 replications) RMSEA values for all (100%) UFA models were $\leq .05$.

RMSEA

RMSEA Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

RMSEA: (M = .04, SD = .03) (M = .05, SD = .03), (M = .08, SD = .05)

Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	All Positive Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	All Negative Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	Mix of Pos/Neg Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)
1000	6	.0	.4	.15	.008	.008	.026	-.018	.008	.008	.025	-.017	.009	.009	.066	-.057
				.30	.007	.007	.029	-.022	.006	.006	.028	-.022	.010	.010	.122	-.113
			.8	.15	.010	.010	.090	-.080	.009	.009	.089	-.080	.010	.010	.162	-.153
				.30	.010	.010	.164	-.154	.009	.009	.163	-.154	.010	.010	.326	-.316
		.5	.4	.15	.006	.006	.014	-.008	.008	.008	.030	-.022	.009	.009	.037	-.028
				.30	.004	.004	.015	-.011	.007	.007	.043	-.036	.009	.009	.065	-.056
			.8	.15	.010	.010	.071	-.061	.009	.009	.076	-.067	.010	.010	.123	-.113
				.30	.010	.010	.158	-.148	.009	.009	.155	-.146	.010	.010	.211	-.201
	24	.0	.4	.15	.004	.004	.018	-.014	.004	.004	.018	-.014	.004	.004	.035	-.031
				.30	.004	.004	.025	-.021	.004	.004	.025	-.021	.004	.004	.072	-.068
			.8	.15	.004	.004	.035	-.031	.004	.004	.035	-.031	.004	.004	.063	-.059
				.30	.004	.004	.069	-.065	.004	.004	.069	-.065	.004	.004	.131	-.127
		.5	.4	.15	.004	.004	.011	-.007	.004	.004	.018	-.014	.004	.004	.026	-.022
				.30	.004	.004	.014	-.010	.004	.004	.033	-.029	.004	.004	.044	-.040
			.8	.15	.004	.004	.030	-.026	.004	.004	.031	-.027	.004	.004	.053	-.049
				.30	.004	.004	.059	-.055	.004	.004	.062	-.058	.004	.004	.110	-.106

Across the 4,320 design conditions, average (across 1,000 replications) RMSEA values for 51% of the CFA models were $\leq$ .05.

RMSEA

RMSEA Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

RMSEA: (M = .04, SD = .03) (M = .05, SD = .03), (M = .08, SD = .05)

Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	All Positive Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	All Negative Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	Mix of Pos/Neg Cross-Loadings: Analysis CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)
1000	6	.0	.4	.15	.008	.008	.026	-.018	.008	.008	.025	-.017	.009	.009	.066	-.057
				.30	.007	.007	.029	-.022	.006	.006	.028	-.022	.010	.010	.122	-.113
			.8	.15	.010	.010	.090	-.080	.009	.009	.089	-.080	.010	.010	.162	-.153
				.30	.010	.010	.164	-.154	.009	.009	.163	-.154	.010	.010	.326	-.316
		.5	.4	.15	.006	.006	.014	-.008	.008	.008	.030	-.022	.009	.009	.037	-.028
				.30	.004	.004	.015	-.011	.007	.007	.043	-.036	.009	.009	.065	-.056
			.8	.15	.010	.010	.071	-.061	.009	.009	.076	-.067	.010	.010	.123	-.113

Many applied researchers might stop here, if they had begun with a CFA model and found an RMSEA ≤ .05

Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)	CFA (CS)	UFA	CFA (IS)	UFA vs. CFA (IS)
		.5	.4	.15	.004	.004	.011	-.007	.004	.004	.018	-.014	.004	.004	.026	-.022
				.30	.004	.004	.014	-.010	.004	.004	.033	-.029	.004	.004	.044	-.040
			.8	.15	.004	.004	.030	-.026	.004	.004	.031	-.027	.004	.004	.053	-.049
				.30	.004	.004	.059	-.055	.004	.004	.062	-.058	.004	.004	.110	-.106

Across the 4,320 design conditions, average (across 1,000 replications) RMSEA values for 51% of the CFA models were ≤ .05.

RMSEA

RMSEA Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

					All Positive Cross-Loadings				All Negative Cross-Loadings				Mix of Pos/Neg Cross-Loadings			
					Analysis			UFA vs.	Analysis			UFA vs.	Analysis			UFA vs.
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)
1000	6	.0	.4	.15	.008	.008	.026	-.018	.008	.008	.025	-.017	.009	.009	.066	-.057
				.30	.007	.007	.029	-.022	.006	.006	.028	-.022	.010	.010	.122	-.113
			.8	.15	.010	.010	.090	-.080	.009	.009	.089	-.080	.010	.010	.162	-.153
				.30	.010	.010	.164	-.154	.009	.009	.163	-.154	.010	.010	.326	-.316
		.5	.4	.15	.006	.006	.014	-.008	.008	.008	.030	-.022	.009	.009	.037	-.028
				.30	.004	.004	.015	-.011	.007	.007	.043	-.036	.009	.009	.065	-.056
			.8	.15	.010	.010	.071	-.061	.009	.009	.076	-.067	.010	.010	.123	-.113
				.30	.010	.010	.158	-.148	.009	.009	.155	-.146	.010	.010	.211	-.201
	24	.0	.4	.15	.004	.004	.018	-.014	.004	.004	.018	-.014	.004	.004	.035	-.031
				.30	.004	.004	.025	-.021	.004	.004	.025	-.021	.004	.004	.072	-.068
			.8	.15	.004	.004	.035	-.031	.004	.004	.035	-.031	.004	.004	.063	-.059
				.30	.004	.004	.069	-.065	.004	.004	.069	-.065	.004	.004	.131	-.127
		.5	.4	.15	.004	.004	.011	-.007	.004	.004	.018	-.014	.004	.004	.026	-.022

If they had not, and decided to conduct UFA anyway, only 15% (of the 4,320 different design conditions) would have found them equivalent (ΔRMSEA $\geq$ -.015)

DISCLAIMER

We do not advocate for the use of cutoffs when gauging the quality of models through GFI and ΔGFI metrics in applied work

We do so as a way of organizing and framing our results.

Results

CFI Results for Selected Conditions with 67% Cross-Loading: All Values Represent the Mean of 1,000 Replications

CFA RMSEA: *M* = .98, *SD* = .01 *M* = .95, *SD* = .03 *M* = .88, *SD* = .09

					All Positive Cross-Loadings				All Negative Cross-Loadings				Mix of Pos/Neg Cross-Loadings			
					Analysis			UFA vs.	Analysis			UFA vs.	Analysis			UFA vs.
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)	CFA (CS)	UFA	CFA (IS)	CFA (IS)
1000	6	.0	.4	.15	.995	.995	.964	.031	.996	.996	.966	.030	.994	.994	.791	.203
				.30	.998	.998	.978	.020	.998	.998	.979	.019	.996	.996	.541	.455
			.8	.15	1.000	1.000	.975	.025	1.000	1.000	.975	.025	1.000	1.000	.919	.081
				.30	1.000	1.000	.935	.065	1.000	1.000	.935	.065	1.000	1.000	.723	.277
		.5	.4	.15	.998	.998	.992	.006	.991	.991	.915	.076	.996	.996	.942	.054
				.30	1.000	.999	.996	.003	.994	.994	.869	.125	.997	.997	.893	.104
			.8	.15	1.000	1.000	.990	.010	.999	.999	.976	.023	1.000	1.000	.958	.042
				.30	1.000	1.000	.977	.023	.999	.999	.882	.117	1.000	1.000	.916	.084

We see similar patterns with the CFI and ΔCFI…

Across the 4,320 design conditions, over 99% of the average (across 1,000 replications) CFI values from UFA models were $\geq$ .95

48% of the CFI values for the CFA models were $\geq$ .95

Only 6% of the UFA vs. CFA contrasts would have found them equivalent (ΔCFI $\leq$ .01)

Results

$RMSEA_D$ Results Comparing UFA with Incorrectly Specified CFA for Selected Conditions: All Values Represent the Mean of 1,000 Replications

$RMSEA_D$ Rejections:					83%			91%			99%		
					67% of Indicators Cross-Load								
					All Positive CLs			All Negative CLs			Mix of Pos/Neg CLs		
						90% CI			90% CI			90% CI	
Sample Size	Num Indic	F-F Cor	TL Mag	CL Mag	*M*	*LB*	*UB*	*M*	*LB*	*UB*	*M*	*LB*	*UB*
1000	6	.0	.4	.15	.042	[.028	, .075]	.041	[.027	, .074]	.094	[.073	, .124]
				.30	.047	[.031	, .079]	.046	[.031	, .079]	.174	[.150	, .202]
			.8	.15	.128	[.105	, .156]	.127	[.105	, .156]	.229	[.205	, .257]
				.30	.232	[.208	, .259]	.231	[.207	, .259]	.462	[.437	, .489]
		.5	.4	.15	.027	[.019	, .062]	.047	[.031	, .079]	.056	[.038	, .087]
				.30	.030	[.020	, .065]	.064	[.045	, .095]	.093	[.072	, .123]
			.8	.15	.102	[.080	, .131]	.109	[.086	, .138]	.174	[.150	, .202]

<u>As well with the $RMSEA_D$…</u>

Across the 4,320 design conditions, 93% of the CFA models would be rejected ($RMSEA_D$ > .05; Sareveli et al., 2023)

Should we be rejecting (CFA) models that constrain minor cross-loadings (< .30; Marsh, et al., 2020) to zero as some suggest for:

CFA evaluations of measurement structure (van Prooijen & van der Kloot, 2001)

(or) calculating factor scores (Grice, 2010)

(or) evaluating simple structure (Tabachnick & Fidell, 2007)

OR,

Should we consider that low cross loadings (< |.30|) when fixed to zero can produce meaningful structural bias (shown here and elsewhere in the literature)?

How might we quantify meaningful structural bias?

Examples of elsewhere: Hsu, et al., 2014; Marsh et al., 2013, 2014; Steenkamp & Maydeu-Olivares, 2023

How about this?

Drawing from recent methodological research in the context of UFA and ESEM, we follow Steenkamp & Maydeu-Olivares (2023) in asserting, on the basis of Cohen's (1988) recommendations, that factor correlation bias $\leq |0.1|$ "are unlikely to be meaningful"

We extend this here for purposes of framing our results around potentially meaningful values, to include bias values $\geq |.30|$ that have been characterized as a 'medium effect' (Cohen, 1988, p. 80) as nonignorable levels of factor correlation bias.

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	. $RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models were found to provide a reasonable fit to the data
- Of these, a little over 20% had ignorable levels of bias (≤ .10)
- And ~ 70% had bias values at or below nonignorable levels (≤ .30)
- However, ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEA [illegible]
"Good Fit"		CFI $\geq$ .95 (48%)	RMSEA $\leq$ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	$\leq$ .10	23%	24%	NA
	$\leq$ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	$\leq$ .10	34%	33%	NA
	$\leq$ .30	64%	66%	NA
CFA ~ UFA		ΔCFI $\leq$ .01 (6%)	ΔRMSEA $\geq$ -.015 (15%)	$RMSEA_D \leq .05$ (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	$\leq$ .10	5%	9%	3%
	$\leq$ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	. $RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	$\leq$ .10	30%	32%	30%
	$\leq$ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models were found to provide a reasonable fit to the data
- Of these, a little over 20% had ignorable levels of bias ($\leq$ .10)
- And ~ 70% had bias values at or below nonignorable levels ($\leq$ .30)
- However, ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	$RMSEA_D$
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models were found to provide a reasonable fit to the data
- Of these, a little over 20% had ignorable levels of bias (≤ .10)
- And ~ 70% had bias values at or below nonignorable levels (≤ .30)
- However, ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	$RMSEA_D$
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models were found to provide a reasonable fit to the data
- Of these, a little over 20% had ignorable levels of bias (≤ .10)
- And ~ 70% had bias values at or below nonignorable levels (≤ .30)
- However, ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	. $RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models had questionable fit to the data
- Of these, a little over 30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And over ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models had questionable fit to the data
- Of these, a little over 30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And over ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	. $RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models had questionable fit to the data
- Of these, a little over 30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And over ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- ~ 50% of the CFA models had questionable fit to the data
- Of these, a little over 30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And over ~ 30% had nonignorable bias values (> .30)

"Good Fit"	Bias	CFI	RMSEA	RMSEAd
		CFI $\geq$ .95 (48%)	RMSEA $\leq$ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	$\leq$ .10	23%	24%	NA
	$\leq$ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	$\leq$ .10	34%	33%	NA
	$\leq$ .30	64%	66%	NA
CFA ~ UFA		ΔCFI < .01 (6%)	ΔRMSEA $\geq$ -.015 (15%)	$RMSEA_D \leq$ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	$\leq$ .10	5%	9%	3%
	$\leq$ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	$\leq$ .10	30%	32%	30%
	$\leq$ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- Far fewer CFA models were deemed similar to their UFA counterparts by way of ΔGFI criteria
- Although a small % of these had had ignorable levels of bias ($\leq$ .10)
- Much larger %'s were observed for bias values at or below nonignorable levels ($\leq$ .30), with the ΔCFI performing best
- However, ~ 30% had nonignorable bias values (> .30)

"Good Fit"	Bias	CFI	RMSEA	RMSEAd
		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI < .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- Far fewer CFA models were deemed similar to their UFA counterparts by way of ΔGFI criteria
- Although a small % of these had had ignorable levels of bias (≤ .10)
- Much larger %'s were observed for bias values at or below nonignorable levels (≤ .30), with the ΔCFI performing best
- However, ~ 30% had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI < .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- Far fewer CFA models were deemed similar to their UFA counterparts by way of ΔGFI criteria
- Although a small % of these had had ignorable levels of bias (≤ .10)
- Much larger %'s were observed for bias values at or below nonignorable levels (≤ .30), with the ΔCFI performing best
- However, ~ 30% had nonignorable bias values (> .30)

"Good Fit"	Bias	CFI	RMSEA	RMSEAd
		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI < .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- Far fewer CFA models were deemed similar to their UFA counterparts by way of ΔGFI criteria
- Although a small % of these had had ignorable levels of bias (≤ .10)
- Much larger %'s were observed for bias values at or below nonignorable levels (≤ .30), with the ΔCFI performing best

However, some ~ 30% had nonignorable bias values (> .30)

Note. ΔCFI did well here

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI $\geq$ .95 (48%)	RMSEA $\leq$ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	$\leq$ .10	23%	24%	NA
	$\leq$ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	$\leq$ .10	34%	33%	NA
	$\leq$ .30	64%	66%	NA
CFA ~ UFA		ΔCFI $\leq$ .01 (6%)	ΔRMSEA $\geq$ -.015 (15%)	$RMSEA_D \leq .05$ (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	$\leq$ .10	5%	9%	3%
	$\leq$ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	$\leq$ .10	30%	32%	30%
	$\leq$ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- The vast majority of CFA models failed equivalence test with their UFA counterparts
- Of the CFA models deemed to provide worse fit ~30% had ignorable levels of bias ($\leq$ .10)
- ~ 65% had bias values at or below nonignorable levels ($\leq$ .30)
- And ~ 30% actually had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- The vast majority of CFA models failed equivalence test with their UFA counterparts
- Of the CFA models deemed to provide worse fit ~30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And ~ 30% actually had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- The vast majority of CFA models failed equivalence test with their UFA counterparts
- Of the CFA models deemed to provide worse fit ~30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And ~ 30% actually had nonignorable bias values (> .30)

	Bias	CFI	RMSEA	RMSEAd
"Good Fit"		CFI ≥ .95 (48%)	RMSEA ≤ .05 (51%)	
	> .10	77%	76%	NA
	> .30	30%	34%	NA
	≤ .10	23%	24%	NA
	≤ .30	70%	67%	NA
"Poor Fit"		CFI < .95 (52%)	RMSEA > .05 (49%)	
	> .10	66%	67%	NA
	> .30	36%	33%	NA
	≤ .10	34%	33%	NA
	≤ .30	64%	66%	NA
CFA ~ UFA		ΔCFI ≤ .01 (6%)	ΔRMSEA ≥ -.015 (15%)	$RMSEA_D$ ≤ .05 (7%)
	> .10	95%	91%	97%
	> .30	10%	32%	27%
	≤ .10	5%	9%	3%
	≤ .30	90%	68%	73%
CFA ≠ CFA		ΔCFI > .01 (94%)	ΔRMSEA < -.015 (85%)	$RMSEA_D$ > 05 (93%)
	> .10	70%	68%	70%
	> .30	35%	33%	34%
	≤ .10	30%	32%	30%
	≤ .30	65%	67%	66%

Based on averages across the 4,320 different design conditions (each with 1,000 replications)

- The vast majority of CFA models failed equivalence test with their UFA counterparts
- Of the CFA models deemed to provide worse fit ~30% had ignorable levels of bias (≤ .10)
- ~ 65% had bias values at or below nonignorable levels (≤ .30)
- And ~ 30% actually had nonignorable bias values (> .30)

CFA models deemed similar to their UFA counterparts by way of ΔGFI criteria were more likely to have the following design conditions:

ΔCFI:

Positive CL condition (75%)…and none of the mixed CL conditions
Lower (33%) CL saturations (76%)
Low (.15) CL values (91%)
High TL ($\geq$ .7) conditions (54%)

ΔRMSEA:

Positive CL condition (55%)…with some mixed CL conditions (8%)
Lower CL saturations (69%)
Low CL values (87%)
Low target loadings (73%)

$RMSEA_D$

Positive CL condition (63%)…with some mixed CL conditions (3%)
Lower CL saturations (75%)
Low CL values (88%)
Low target loadings (66%)
Smaller models (50%)

Partial Eta-Squared Values for Raw Factor-Factor Correlation Bias and UFA vs. CFA GFI Change by Population Design Condition

Population Design Conditions	Raw FF R Bias	Δ CFI	Δ RMSEA	$RMSEA_D$
Main Effects				
N	0.00	0.04	0.71	0.04
Model Size	0.01	0.11	0.96	0.73
Factor Correlation (FC)	0.21	0.61	0.75	0.88
Target Loadings (TL)	0.20	0.61	0.98	0.99
Cross Loading (CL) Value	0.20	0.92	0.98	0.99
CL Sign	1.00	0.96	0.98	0.99
CL Saturation	0.01	0.74	0.91	0.97
Two-Way Interactions				
N * Model Size	0.00	0.01	0.11	0.00
N * FC	0.00	0.00	0.01	0.01
N * TL	0.00	0.02	0.08	0.03
N * CL Value	0.00	0.01	0.01	0.00
N * CL Sign	0.00	0.04	0.06	0.05
N * CL Saturation	0.00	0.00	0.00	0.00
Model Size * FC	0.01	0.08	0.40	0.08
Model Size * TL	0.02	0.12	0.91	0.61
Model Size * CL Value	0.03	0.01	0.73	0.51
Model Size * CL Sign	0.04	0.05	0.81	0.47
Model Size * CL Saturation	0.00	0.00	0.43	0.20
FC * TL	0.26	0.08	0.16	0.24
FC * CL Value	0.12	0.27	0.24	0.44
FC * CL Sign	0.85	0.84	0.72	0.85
FC * CL Saturation	0.07	0.06	0.05	0.12
TL * CL Value	0.11	0.05	0.87	0.94
TL * CL Sign	0.92	0.63	0.80	0.91
TL * CL Saturation	0.04	0.01	0.67	0.83
CL Value * CL Sign	0.96	0.88	0.87	0.95
CL Value * CL Saturation	0.00	0.25	0.38	0.64
CL Sign * CL Saturation	0.95	0.70	0.70	0.86

Population Design Conditions	Raw FF R Bias	Δ CFI	Δ RMSEA	$RMSEA_D$
Main Effects				
N	0.00	0.04	0.71	0.04
Model Size	0.01	0.11	0.96	0.73
Factor Correlation (FC)	0.21	0.61	0.75	0.88
Target Loadings (TL)	0.20	0.61	0.98	0.99
Cross Loading (CL) Value	0.20	0.92	0.98	0.99
CL Sign	1.00	0.96	0.98	0.99
CL Saturation	0.01	0.74	0.91	0.97
Two-Way Interactions				
N * Model Size	0.00	0.01	0.11	0.00
N * FC	0.00	0.00	0.01	0.01
N * TL	0.00	0.02	0.08	0.03
N * CL Value	0.00	0.01	0.01	0.00
N * CL Sign	0.00	0.04	0.06	0.05
N * CL Saturation	0.00	0.00	0.00	0.00
Model Size * FC	0.01	0.08	0.40	0.08
Model Size * TL	0.02	0.12	0.91	0.61
Model Size * CL Value	0.03	0.01	0.73	0.51
Model Size * CL Sign	0.04	0.05	0.81	0.47
Model Size * CL Saturation	0.00	0.00	0.43	0.20
FC * TL	0.26	0.08	0.16	0.24
FC * CL Value	0.12	0.27	0.24	0.44
FC * CL Sign	0.85	0.84	0.72	0.85
FC * CL Saturation	0.07	0.06	0.05	0.12
TL * CL Value	0.11	0.05	0.87	0.94
TL * CL Sign	0.92	0.63	0.80	0.91
TL * CL Saturation	0.04	0.01	0.67	0.83
CL Value * CL Sign	0.96	0.88	0.87	0.95
CL Value * CL Saturation	0.00	0.25	0.38	0.64
CL Sign * CL Saturation	0.95	0.70	0.70	0.86

Partial Eta-Squared Values for Raw Factor-Factor Correlation Bias and UFA vs. CFA GFI Change by Population Design Condition

Closing Thoughts

- GFI measures provide a convenient tool for gauging model fit, however, there seems to be many false negatives and false positives (in relation to FF correlation bias) when used to evaluate CFA models (relative to UFA models)

- It may have come for us to reconsider what constitutes a 'trivial misspecification' and to more fully embrace the idea that unidimensional indicators are "an inconvenient fiction" (Marsh et al., 2013, p. 258)

- We encourage users of CFA to not solely rely on stand-alone GFI measures, even when they are 'good' in relation to historical thresholds, but to also examine and consider modification indices that might point to misspecifications in CFA models that could lead to less biased parameter estimates among structural components

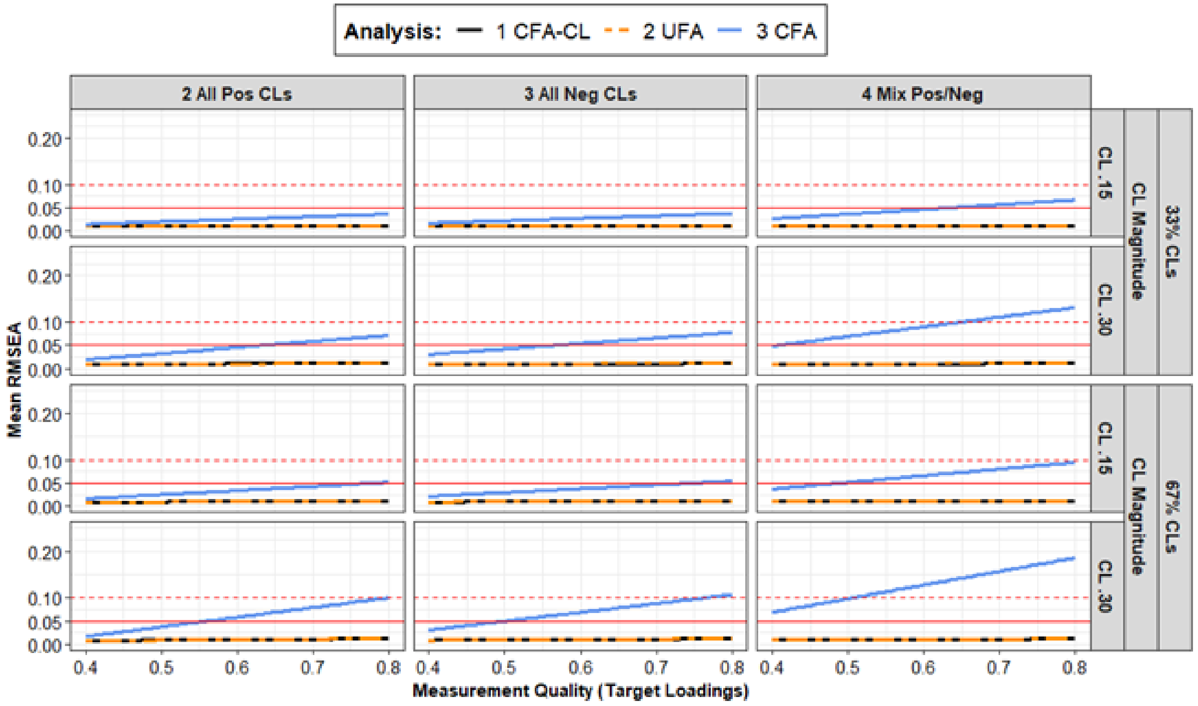
- We encourage users of UFA/ESEM to also include examination their estimated model parameters when evaluating their models, regardless of "strong" GFI values that may accompany their models.
    - In the current study, nearly all UFA models produced GFI values suggestive of good fit.
    - We are not particularly pleased with the amount of factor correlation bias present in the UFA models.
    - This has been shown elsewhere (e.g., Morin et al., 2013).
    - There remains room for improvement in UFA bias that might come from inspection model results that reveal (along with theory) that some aspects of the UFA measurement model could be constrained to zero

Appendix

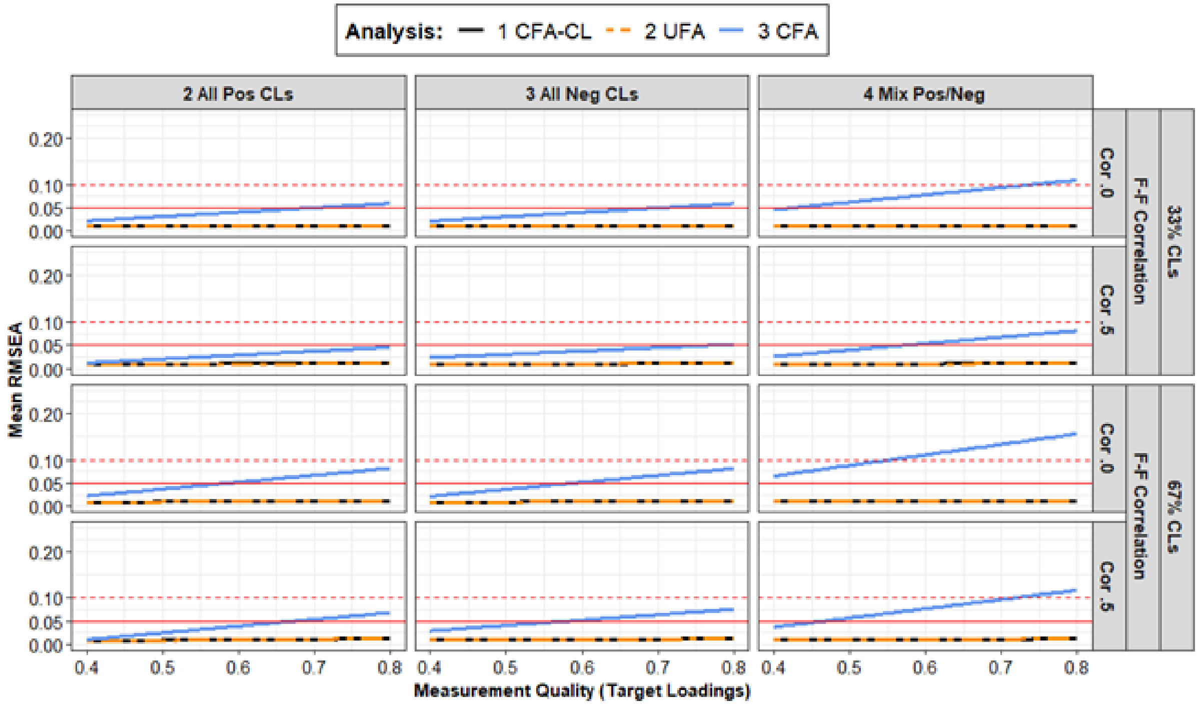
Results

Results for Mean RMSEA Values by Analysis

CL Magnitude



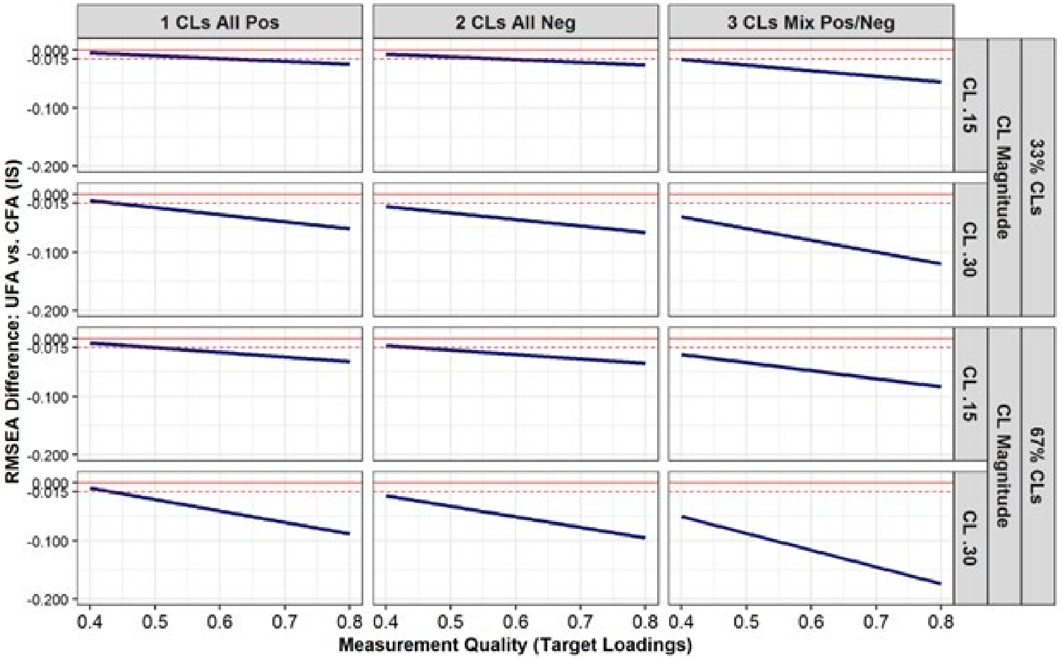
F-F Correlation



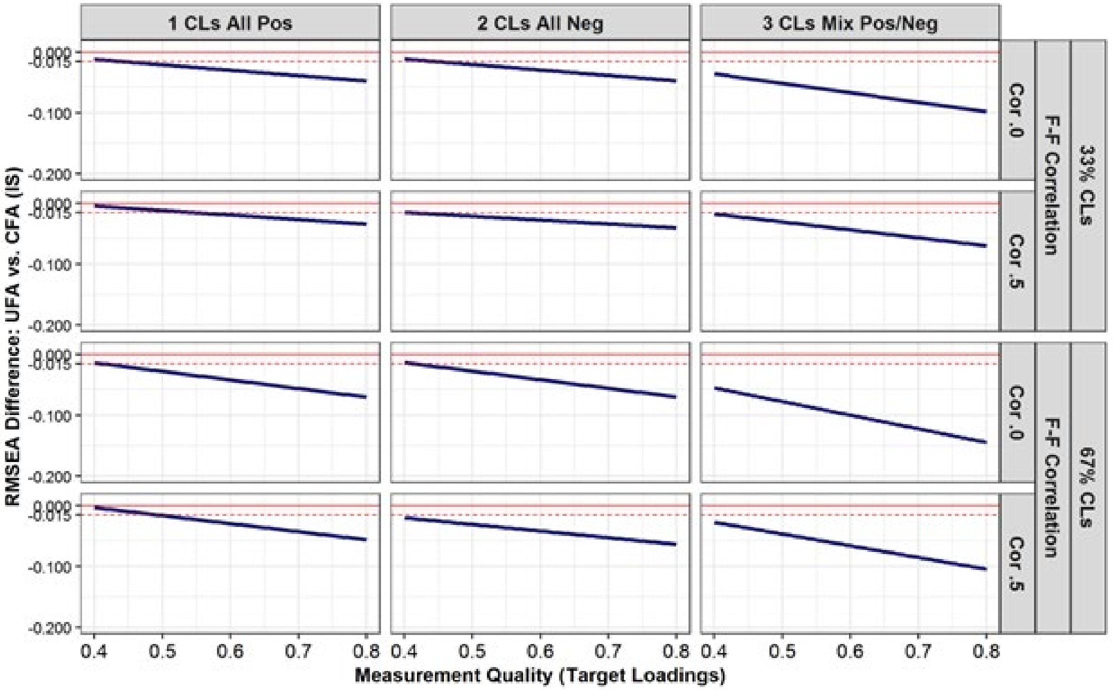
Results

Results for RMSEA Differences between UFA and CFA (IS)

CL Magnitude



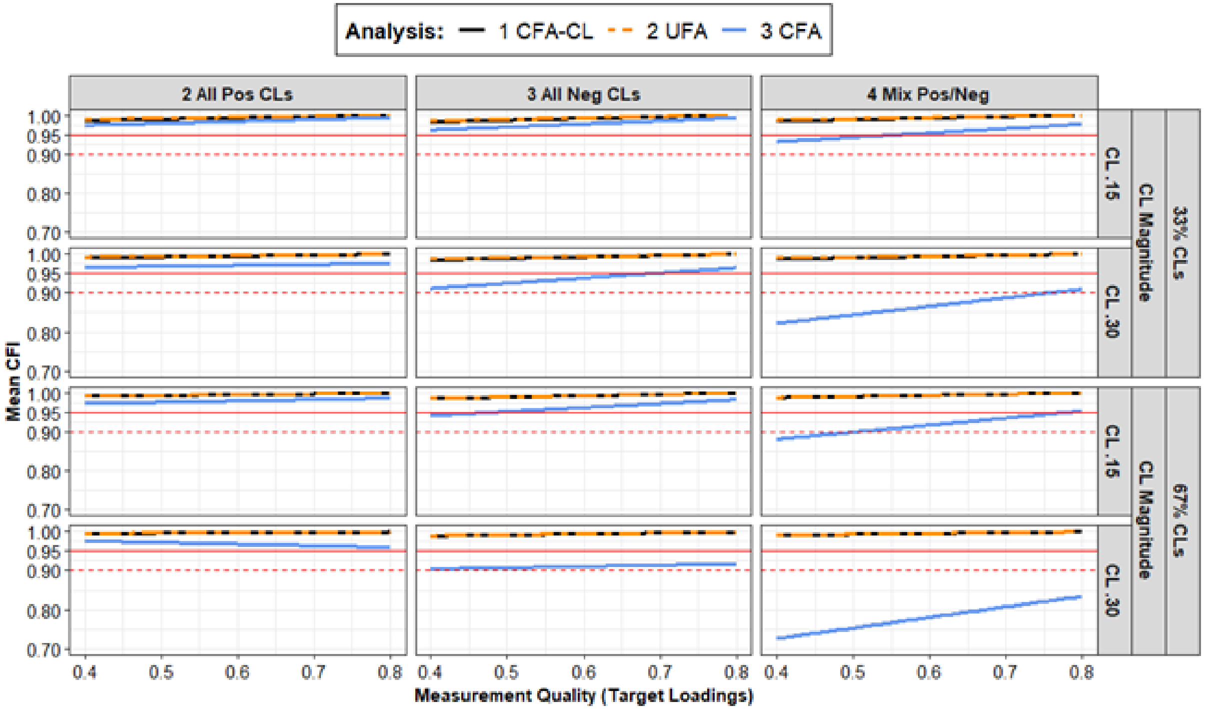
F-F Correlation



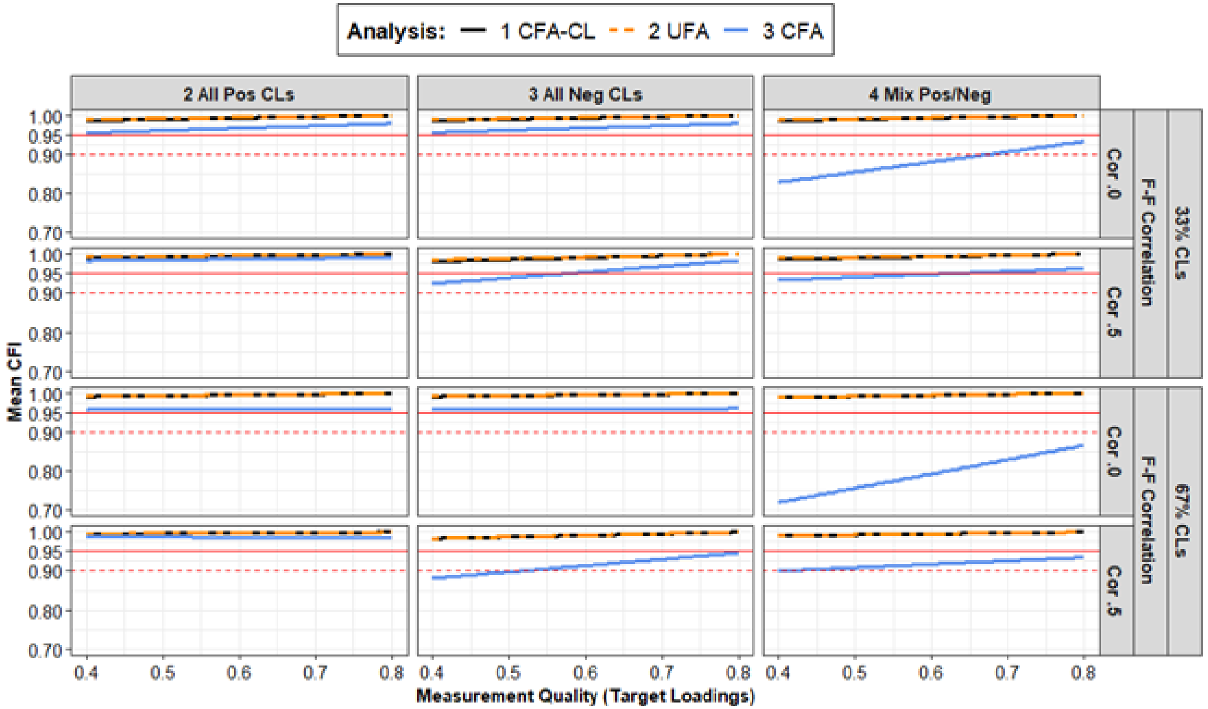
Results

Results for Mean CFI Values by Analysis

CL Magnitude



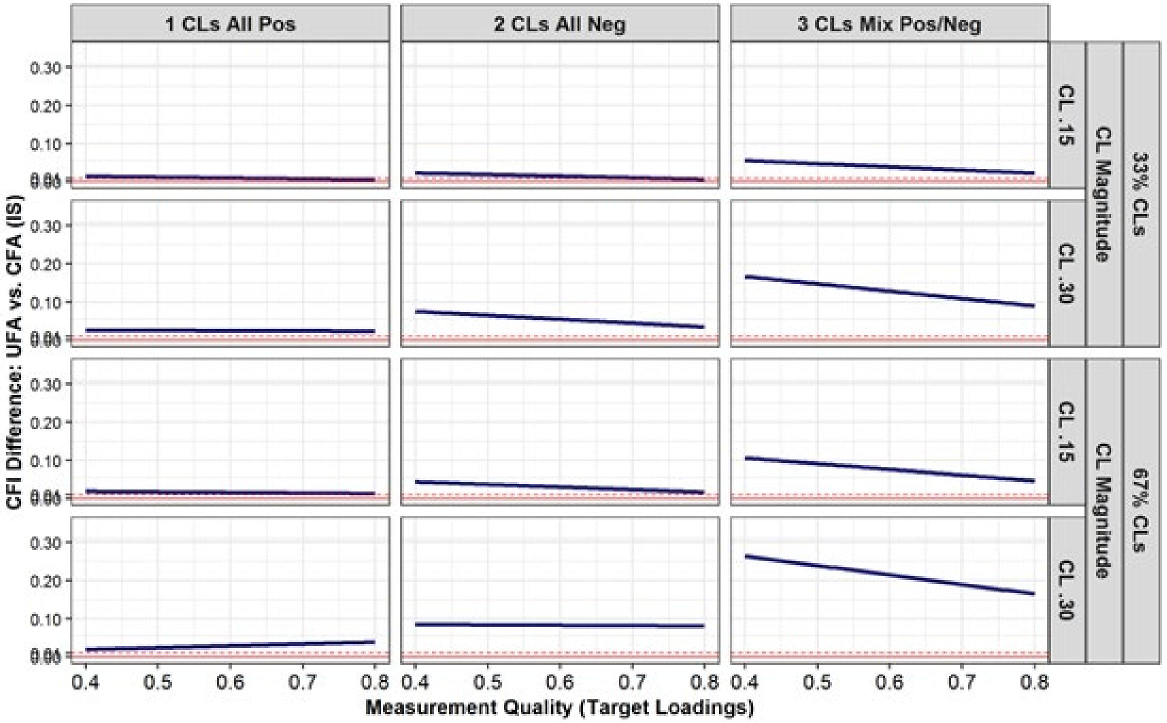
F-F Correlation



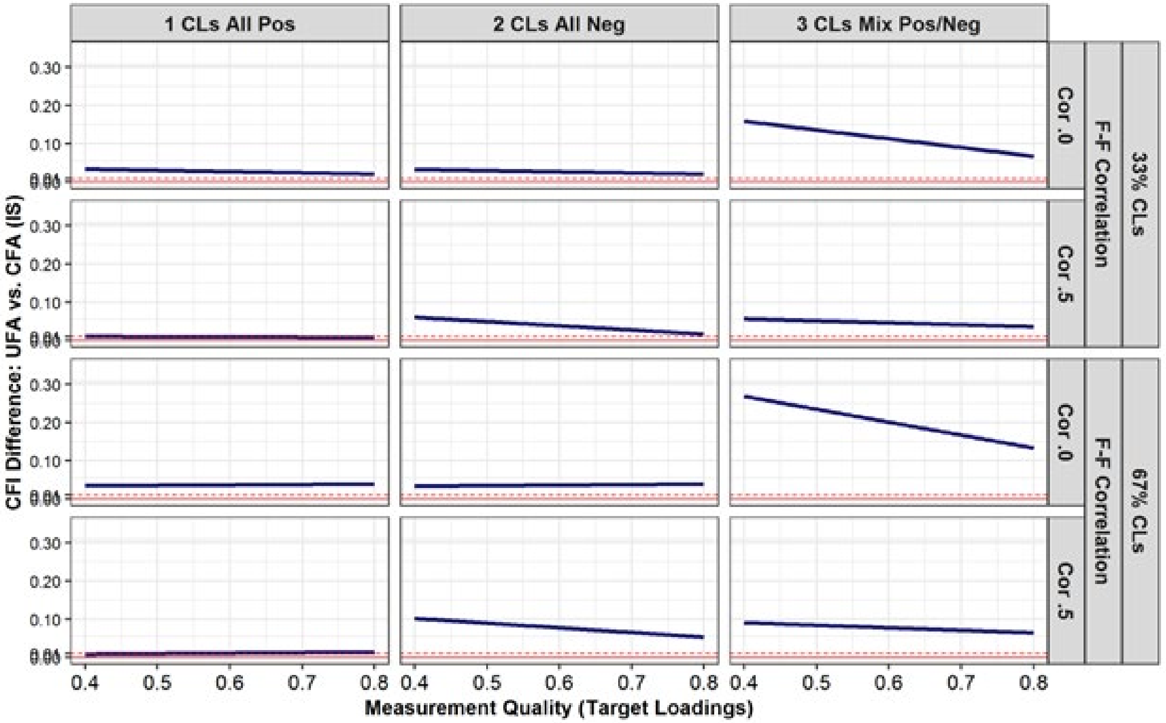
Results

Results for CFI Differences between UFA and CFA (IS)

CL Magnitude



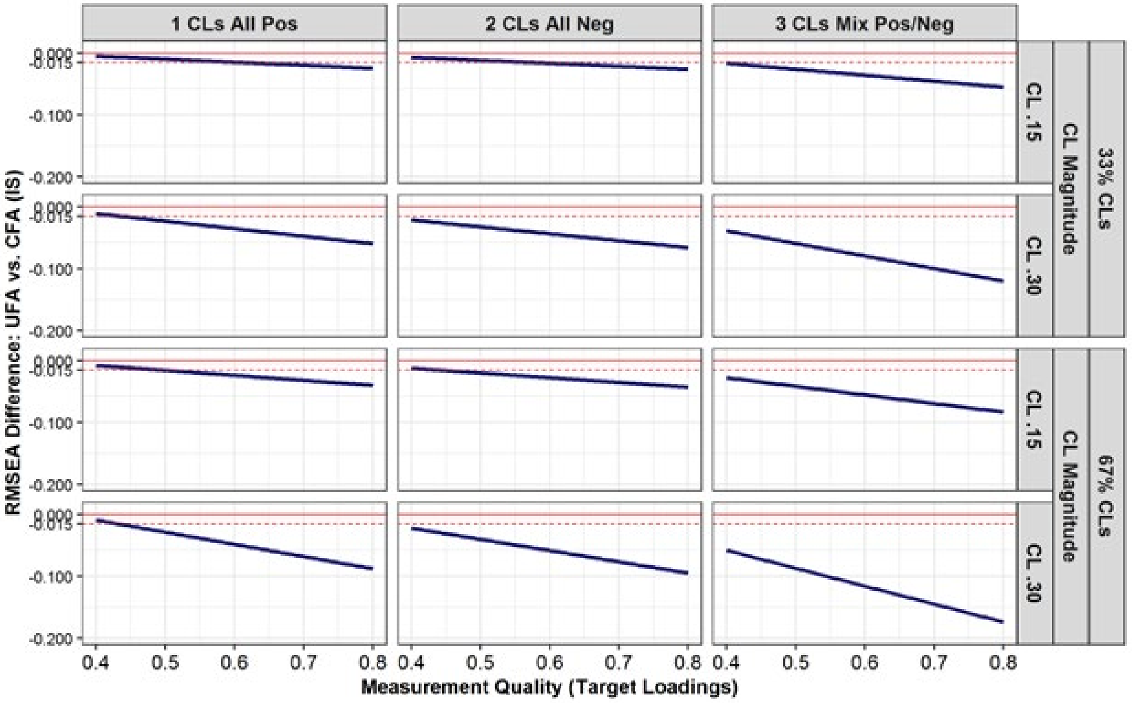
F-F Correlation



Results

Results for RMSEA Differences between UFA and CFA (IS)

CL Magnitude



F-F Correlation