Meade, A. W. (2005, April). *Sample Size and Tests of Measurement Invariance*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Sample Size and Tests of Measurement Invariance

Adam W. Meade
North Carolina State University

Though widely used, confirmatory factor analysis tests of measurement invariance are not well understood. Results of a simulation study indicated that the power of invariance tests varied widely depending on sample size, factor over-determination, and item communality. Accurate estimation of parameters provide a possible explanation for these results.

Measurement equivalence/invariance (ME/I) can be thought of as operations yielding measures of the same attribute under different conditions (Horn & McArdle, 1992). These different conditions include stability of measurement over time (Golembiewski, Billingsley, & Yeager, 1975), across different populations (e.g., cultures; Riordan & Vandenberg, 1994), different mediums of measurement administration (e.g., web-based survey administration verses paper and pencil measures; Taris, Bok, & Meijer, 1998), or different rater groups in multi-source feedback assessments (Facteau & Craig, 2001; Vandenberg, 2002). Under each of these conditions, tests of ME/I are often conducted via confirmatory factor analytic (CFA) methods. These methods have evolved substantially during the past twenty years and are in widespread use (Vandenberg & Lance, 2000). Recently, however, Vandenberg (2002) has made a call for increased research on ME/I analyses and the logic behind them, stating that "A negative aspect of this fervor, however, is characterized by unquestioning faith on the part of some that the technique is correct or valid under all circumstances" (p. 140).

Although hundreds of studies using CFA tests of ME/I have been conducted, as of yet researchers have not fully delineated the mechanism via which these tests assess invariance. Using MacCallum, Widaman, Zhang, & Hong's (1999) seminal study as a guide, this study illustrates the effects of sample size, item communality, and factor over-determination on the power of CFA tests of equality of factor loadings. Moreover, we provide a possible explanatory mechanism for these effects.

CFA tests of ME/I Overview

CFA tests of ME/I involve simultaneously fitting a series of measurement models to two or more data samples. Typically, in the first model, both data sets (representing groups, time periods, etc.) are examined simultaneously, holding only the pattern of factor loadings invariant. This baseline model of equal factor patterns provides a chi-square value that reflects model fit for item parameters estimated independently yet simultaneously for each group. Next, a test of factor loading invariance (metric invariance; Horn & McArdle, 1992) is conducted by examining a model identical to the baseline model except that the matrix of factor loadings (Λ_x) is constrained to be equal across groups. A significant change in chi-square indicates that factor loadings differ (i.e., items do not relate to the factors in the same way) across groups.

Subsequent to tests of factor loadings, several additional model parameters may be tested for ME/I (see Vandenberg & Lance, 2000, for a review). However, in this study we chose to focus on the effects of data properties on ME/I analyses of factor loadings. We choose to limit our investigation to these parameters because, presuming the same general model holds in both samples, there is firm consensus that tests of factor loadings are the most important for establishing that ME/I conditions exist (Vandenberg & Lance, 2000).

While there are many examples of applications of tests of ME/I in the extant literature (see Vandenberg & Lance, 2000; Riordan, Richardson, Schaffer, & Vandenberg, 2001), little has been done to determine the factors that affect the efficacy of these tests (Meade & Lautenschlager, 2004; Vandenberg, 2002). In one published study, Meade and Lautenschlager (2004) manipulated the number of items showing psychometric differences (sometimes called differential functioning, DF, items) between simulated groups, and the directionality of the simulated differences. They found that a lack of ME/I was more readily detected when more items were simulated to differ. Moreover, they found less frequent detection of a lack of ME/I when factor loadings were always higher in one group than when some were higher in both groups.

The Current Study

This study attempts to provide an explanatory mechanism for the efficacy of CFA tests of equality of factor loadings. We posit that tests of factor loadings constitute tests of equality of estimated parameters. As such, larger and more numerous differences in the estimated factor loadings for the two groups will result in an increased probability that the metric invariance test will be significant (Meade & Lautenschlager, 2004). However, other properties of the data can affect the efficacy of detecting these differences.

We propose a two part explanation of the effects of data properties on the detection DF items. First, data properties (such as sample size, factor over-determination, and item communality) directly affect the accuracy of parameter estimation in the CFA model. Secondly, this accuracy of estimation is directly related to the probability that the test of metric invariance will be significant. If accuracy of parameter detection is directly related to the probability that metric invariance tests will be significant, there is a need to delineate the variables which affect the accuracy of estimation of factor loadings. In our study, we simulated a number of conditions of DF items (i.e., a lack of ME/I) between a pair of groups. In each condition, we held the population magnitude of differences between groups constant, but varied other data properties. In order to guide our search for data properties that affect the accuracy of parameter estimation, we relied primarily on the work of MacCallum et al. (1999). MacCallum and colleagues noted that within exploratory factor analysis, greater factor over-determination (having more indictors per factor), higher item communality, and larger sample size led to more accurate estimated factor loadings. We expanded the work of MacCallum et al. (1999) to the two group case of CFA tests of ME/I to determine the effect of these variables on CFA tests of metric invariance.

Method

In this study, we simulated data for two groups of respondents. Data for hypothetical Group 1 respondents was simulated, then Group 1 data properties were manipulated to create Group 2 data.

We based our simulation conditions in part based on the values chosen by MacCallum et al. (1999) for sample size, factor over-determination, and item communality. The factor over-determination conditions required simulating separate Group 1 data sets with different numbers of latent factors. We choose to simulate a condition of low factor over-determination in which twenty items represent six orthogonal factors and a high factor over-determination condition in which twenty items represent three orthogonal factors. Item communalities ranged from .2 to .7 within each factor (similar to MacCallum et al.'s wide communality condition) in both six and three factor data.

Data Properties Manipulated

For all conditions in this study, DF of items was simulated by changing the magnitude of factor loadings for four Group 1 items by .20 to create Group 2 factor loadings (see Meade & Lautenschlager, 2004).

Sample size. In this study, there were three conditions of sample size, *N*=100, 200 and 400 per group of respondents. These values correspond closely to those chosen by MacCallum et al., and initial pilot analyses suggested that these values tended to adequately illustrate the effects of sample size.

Item communality. Item communalities were manipulated in two primary ways: By choice of which items exhibited DF (either low or high communality items), or the manner in which the items were modified (in uniform or mixed fashion) to create Group 2 item parameters.

The choice of items to modify (thus creating DF items) was either low or high communality items. For the low communality choice condition, the four items with the lowest Group 1 factor loadings were modified to create the Group 2 factor loadings. For the high communality DF item condition, the four highest Group 1 factor loadings were modified to create the Group 2 factor loadings. In all cases, factor loadings for DF items were modified by the same magnitude, .2. However, as item communality for these data is equal to the square of the factor loadings, a change of .2 for a higher factor loading constitutes a larger change in communality than does an equal change in a lower factor loading.

For the *how modified* variable, Group 1 factor loadings were either modified in a uniform or mixed pattern. For the uniform pattern condition, a value of .2 was subtracted from Group 1 factor loadings for all four DF items to create Group 2 factor loadings. For the mixed pattern condition, a value of .2 was *subtracted* from two Group 1 DF items while a value of .2 was *added* to two other DF items. See Table 1 for an overview of the simulation conditions. There were 500 replications for all conditions in this study.

Analyses

For each of the 500 replications in each condition, tests of metric invariance were conducted by examining the difference in chi-square values between a baseline model, in which factor loadings were estimated simultaneously yet were allowed to vary between groups, and a model in which all item parameters were constrained to be equal for the groups.

Accuracy of parameter recovery was assessed by examining the root mean squared error (RMSE) between the population simulated parameters and those recovered during the analyses. As the effects of accuracy of estimation of factor loadings on metric invariance tests applies primarily to the four items that differ between groups, we computed RMSE as the square root of the mean of the four Group 1 and Group 2 DF items' squared difference between the population and estimated parameters.

Results

Results are presented in three parts. First, the effects of sample size, factor over-determination, and item communality on accuracy of factor loading estimation are presented. Next, differences in parameter estimation accuracy are presented for significant and non-significant metric invariance tests. Lastly, the effects of the sample properties on the significance of tests of metric invariance are presented.

Results of the effects of study variables on DF item RMSE values are presented in Table 2. Table 3 presents the RMSE values by study condition while Figure 1 charts these same values. Overall, the study variables accounted for 41% of the variance in RMSE values ($F_{23,11626}$ = 349.31, $p < .01$). As can be seen in Table 2, all four study variables had a main effect on RMSE values, while 5 of the 6 two-way interactions were significant, as well as the three-way interaction of sample size, number of factors, and which items were chosen as DF items. However, it appears that only the three-way interaction and two of the two-way interactions had noticeable effect sizes.

Examination of Figure 1 and Table 3 aide in interpretation of these effects. While main effects and two-way interactions are generally not directly interpretable in the presence of higher-order interactions, the directionality of two main effects and the two-way interactions in this study were maintained across all levels of the other interactive variables. As can be seen in Figure 1, all other things held constant, larger sample sizes were always associated with smaller RMSE values than smaller sample sizes, while higher factor over-determination (i.e., 3 factor) data had lower RMSE values than lower factor over-determination conditions. A sizable two-way interaction was found between sample size and the number of factors where three factor data had lower RMSE values than six factor data, however, these differences were more pronounced for smaller sample sizes.

The interaction between the number of factors and which items were chosen to be DF items was more complex. The effects of choice of DF items on RMSE values had opposite effects for different levels of number of factors. Specifically, RMSE values were lower when high communality items were modified than when low communality items were modified when there were three factors (i.e., high over-determination) while the opposite pattern emerged for six factors (i.e., low over-determination). The significant three way interaction included these same two variables plus sample size. As can be seen in Figure 1, the interaction between choice of DF item and number of factors varied across sample sizes. The interactive effect was minimal for sample sizes of 400 but more pronounced for smaller sample sizes.

While the mean difference between Group 1 and Group 2 estimated DF factor loadings did vary by condition, these differences were trivial, centering around their population value of .20. More informative is the standard deviation of this difference across the 500 replications in the study (see Table 4). Like the RMSE, this standard deviation indicates the accuracy with which the factor loadings were estimated for the four DF items in the two groups. Analyses of these effects were virtually identical to those of RMSE values. Comparing Figures 1 and Figures 2 highlights the similarity of these effects.

Of all metric invariance tests, 3859 tests of were significant while 7791 were not[1]. As expected, the mean RMSE values were significantly lower for significant metric invariance tests (M=.044, SD=.053) than non-significant metric invariance tests (M=.078, SD=.021; $F_{1,11684}$=1468.55; R^2=.11). Moreover, this trend held after controlling for all other study variables, indicating that even within a particular study condition, significant metric invariance tests tended to have more accurately estimated DF items' factor loadings for than did non-significant metric invariance tests.

Lastly, results of the metric invariance tests were regressed onto the study variables. We created dummy variables and used logistic regression to

[1] Approximately 350 analyses failed to yield proper solutions and were excluded from further analyses.

examine the effects of the study variables on the dichotomous significant/non-significant metric invariance test dependent variable. The overall logistic regression model was significant ($Wald_{23}$ = 2615.47, $p < .01$). While no directly comparable index of R^2 is available in logistic regression, the Cox and Snell's approximation to R^2 was .50 while the rescaled Nagelkerke R^2 static was .70[2] (see Hair, Anderson, Tatham, & Black, 1998 for a review). Taken together, these indices indicate that a large proportion of variance in the significance of metric invariance tests was due to the study variables. Wald significance statistics, standardized parameter estimates, odds-ratios, and their associated confidence intervals for individual study variables can be found in Table 5.

As can be seen in Table 5, results of these analyses were similar to those of the RMSE values. Main effects for all study variables were found as well as three two-way interactions and a three-way interaction. Examination of the standardized coefficients and odds-ratios indicate that, as before, the two-way interaction between which items were DF items and the number of factors was particularly strong. Additionally, the interaction between these two variables was more pronounced for large sample sizes and was virtually non-existent for sample sizes of 100 resulting in a significant three-way interaction.

The two-way interaction of which items were DF items and how communality was manipulated was also significant. Examination of Table 6 and Figure 3 indicates that this interaction was due to a large effect of how the communality was manipulated for three factor data. Unlike RMSE analyses, metric invariance tests were significant much more often for high communality DF items under the mixed patterns of changes condition than they were under the uniformly lower changes condition for three item data (particularly for N=200). Lastly, though the two-way interaction of sample size and the number of factors was significant, these results were contingent upon which items were chosen as DF items, as indicated by the three way-interaction of these variables.

As with RMSE values, although higher order interactions were significant, holding other variables constant, larger sample sizes and a mixed pattern of changes in communality were always associated with a larger number significant metric invariance tests. However, the magnitude of these differences varied greatly depending on the levels of other study variables.

[2] The maximum value of the Cox and Snell R^2 statistic is less than 1.0 while the Nagelkerke R^2 static is rescaled to have a maximum value of 1.0.

Discussion

Perhaps the most important finding from this study is that the data properties, not just sample size, must be considered when determining power for ME/I tests. While sample size did have perhaps the strongest effect on the percentage of metric invariance tests that were significant, by no means were the effects of other study variables trivial. Our results indicated that under all conditions simulated, power was low for sample sizes of 100, while power differed considerably by condition for sample sizes of 200 and 400. For sample sizes of 200 per group, the percentage of metric invariance tests that were significant varied from 4% to 83% and for sample sizes of 400 per group, varied from 50% to 100%. The strong effects of sample size are not surprising given that sample size is directly incorporated into the formula for computing the chi-square statistic used in metric invariance tests. Thus sample size affects both the chi-square statistic directly and the precision of estimation of factor loading parameters.

A second contribution of this study was an explanatory mechanism for the results found. Our analyses for the effects of study variables on the precision of estimation of factor loadings largely mirrored those of the results of the metric invariance tests. Moreover, significant metric invariance tests had more precise estimates of factor loadings of DF items than did non-significant metric invariance tests, even when controlling for study variables. These results confirm that the results of metric invariance tests are closely associated with accurate estimation of DF item parameters. The precision of the estimated factor loadings is, in turn, dependent upon sample size, over-identification of the factors, and item communalities.

While several of the interactions between study variables were significant, results for sample size and the manner in which DF items differed were directly interpretable main effects. Holding other study variables constant, larger sample sizes and a mixed pattern of changes in factor loadings were always associated with both a larger number of significant metric invariance tests and more accurate DF item factor loading estimates than were smaller sample sizes and uniformly lower factor loading changes. These results parallel those from Meade and Lautenschlager (2004). However, as evidenced by the interactions of the study variables, the magnitude of these differences varied greatly by the amount of factor over-identification and which items were chosen as DF items.

One consistent finding in this study was the interaction between factor over-identification and the choice of which items were DF items. For the case of high levels of over-identification (i.e., three factors and twenty items), when DF items communalities were high, factor loadings were more accurately estimated and metric invariance tests were more likely to be significant than when DF items communalities were low. However, when factor over-identification was low (i.e., six factors and twenty items), factor loading were more accurately estimated and metric invariance tests were more likely to be significant when DF items communalities were low. For these data item communality was equal to the square of the factor loadings, thus factor loading changes of .2 for items with large factor loadings resulted in a greater change in communality than changes to items with small factor loadings. We believe that in the six factor case, larger changes in communality lead to a more imprecision of estimation of the factor loadings (as was found) due to the larger role of error variance in the common factor than in the three factor case. This imprecision subsequently results in fewer significant metric invariance tests. For the three factor case, there are enough indicators that error plays a lesser role in the variance of the common factor, thus larger changes in communality are more easily detected than are smaller changes.

Limitations & Future Research

Even though this was a large simulation study, this study was still limited by scope. There are an infinite number of ways in which data can vary and we considered only a very small subset of these possibilities. As a result, most researchers would be unlikely to reference this study as a justification of adequate power to detect a lack of ME/I with their own data. However, the goal of this study was not to provide specific estimates of power for all configurations of data properties. Instead, we sought to provide an initial exploration of the general data properties that affect power in ME/I tests. Moreover, we sought to provide an explanatory mechanism for the effects of these data properties, namely precision of estimation of population differences between groups. Another limitation is that we only considered tests of metric invariance. While these test are typically considered to be the most important ME/I test (presuming the same general factor structure holds in both samples), the effects of data properties on other ME/I tests would be an additional substantial contribution.

This study illustrates the complex relationships between data properties and the efficacy of metric invariance ME/I tests. Though only a few conditions were simulated, we hope that future researchers can build upon this study in further delineating the precise role of factor over-determination, item communalities, and sample size in order to provide mathematical rules of power calculation for ME/I tests.

References

Facteau, J. D., & Craig, S. B. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology, 86*(2), 215-227.

Golembiewski, R. T., Billingsley, K., & Yeager, S. (1975). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science, 12*(2), 133-157.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3-4), 117-144.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: Users Reference Guide.* Chicago: Scientific Software International.

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

Meade, A. W., & Lautenschlager, G. J. (2004). A Monte-Carlo Study of Confirmatory Factor Analytic Tests of Measurement Equivalence/Invariance. *Structural Equation Modeling, 11*(1), 60-72.

Riordan, C. M., Richardson, H. A., Schaffer, B. S., & Vandenberg, R. J. (2001). Alpha, beta, and gamma change: A review of past research with recommendations for new directions. In C. A. Schriesheim & L. L. Neider (Eds.), *Equivalence of Measurement.* Greenwich, CT:: Information Age Publishing.

Riordan, C. M., & Vandenberg, R. J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management, 20*(3), 643-671.

Taris, T. W., Bok, I. A., & Meijer, Z. Y. (1998). Assessing stability and change of psychometric properties of multi-item concepts across different situations: A general approach. *Journal of Psychology, 132*(3), 301-316.

Vandenberg, R. J. (2002). Toward a further understanding of an improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*(2), 139-158.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-69.

Author Contact Info:

Adam W. Meade
Department of Psychology
North Carolina State University
Campus Box 7801
Raleigh, NC 27695-7801
Phone: 919-513-4857
Fax: 919-515-1716
E-mail: adam_meade@ncsu.edu

Table 1.

Overview of Study Conditions

Group 1 Data Properties	
Number of Items	20
Number of Factors	3, 6
Factor Loading Values	Range from .45 to .83

Study Variables used to Create Group 2 Data from Group 1 Data

Variables	Conditions
Sample Size	100, 200, 400 per group
Number of Factors	3, 6
Communality Manipulations:	
Which Items Manipulated	Low or High Communality Items
How Items Manipulated	Uniformly Lower Group 2 Communality or Mixed Pattern of Change of Communality for Group 2

Note: Group 1 and Group 2 data always contained same sample size and same number of factors.

Table 2.

Effects of study variables on RMSE values

Source	*df*	*F*	η^2
Sample Size (N)	2	2181.50**	0.22
# Factors (F)	1	1987.28**	0.10
Which item manipulated (W)	1	25.64**	0.00
How items manipulated (H)	1	128.60**	0.01
F*W	1	810.31**	0.04
N*F	2	263.64**	0.03
N*H	2	3.84*	0.00
N*W	2	38.91**	0.00
F*H	1	1.79	0.00
W*H	1	4.72*	0.00
N*F*W	2	150.69**	0.02
N*F*H	2	0.31	0.00
N*W*H	2	0.63	0.00
F*W*H	1	2.45	0.00
N*F*W*H	2	0.07	0.00

Note: * p<.05, ** p<.01.

Table 3.

RMSE values for Study Conditions

	K=3				K=6			
	All Lower		Mixed		All Lower		Mixed	
N	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
100	0.091	0.064	0.083	0.049	0.106	0.149	0.099	0.139
200	0.064	0.045	0.058	0.034	0.076	0.082	0.069	0.075
400	0.046	0.031	0.042	0.024	0.052	0.052	0.047	0.048

Note: K = number of factors, All lower and Mixed indicate the manner with which factor loadings were manipulated, Lowest and Highest indicate which factor loadings were manipulated.

Table 4.

Standard Deviation of Difference between Group 1 and Group 2 Factor Loading Estimates for Four DF Items

	K=3				K=6			
	All Lower		Mixed		All Lower		Mixed	
N	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
100	0.051	0.038	0.047	0.033	0.060	0.105	0.060	0.105
200	0.036	0.027	0.033	0.024	0.048	0.051	0.043	0.050
400	0.027	0.019	0.024	0.016	0.033	0.032	0.029	0.029

Note: K = number of factors, All lower and Mixed indicate the manner with which factor loadings were manipulated, Lowest and Highest indicate which factor loadings were manipulated.

Table 5.

Results of Logistic Regression of Metric Invariance Test on Study Variables

Parameter	β	SE	Wald Statistic	Odds Ratio	Odds Ratio 95% Confidence Interval	
Intercept	3.49	0.40	77.24**			
N100	-2.85	0.36	63.45**	0.00	<0.01	0.01
N200	-1.32	0.09	224.20**	0.07	0.05	0.10
# Factors (F)	0.23	0.07	11.11**	1.58	1.21	2.08
Which (W)	-0.29	0.06	20.15**	0.56	0.43	0.72
How (H)	0.17	0.07	6.48*	1.42	1.08	1.85
N100*H	0.23	0.93	0.06	1.26	0.21	7.74
N200*H	0.18	0.24	0.53	1.19	0.74	1.92
N100*W	-0.31	1.01	0.10	0.73	0.10	5.30
N200*W	0.51	0.30	2.81	1.66	0.92	3.00
N100*F	12.69	263.90	0.00	>999.99	<0.01	>999.99
N200*F	0.85	0.26	10.37**	2.34	1.39	3.91
W*F	-4.15	0.48	75.22**	0.02	0.01	0.04
W*H	-0.03	0.19	0.03	0.97	0.67	1.40
F*H	-0.72	0.22	10.48**	0.49	0.32	0.75
N100*H*F	-11.01	263.90	0.00	<0.01	<0.01	>999.99
N200*H*F	0.47	0.37	1.60	1.60	0.77	3.31
N100*H*W	0.21	1.37	0.02	1.23	0.08	18.02
N200*H*W	0.16	0.42	0.15	1.17	0.52	2.68
N100*F*W	-8.37	263.90	0.00	<0.01	<0.01	>999.99
N200*F*W	1.71	0.59	8.58**	5.55	1.76	17.46
H*F*W	-11.77	277.50	0.00	<0.01	<0.01	>999.99
N100*H*F*W	20.96	383.00	0.00	>999.99	<0.01	>999.99
N200*H*F*W	9.36	277.50	0.00	>999.99	<0.01	>999.99

Note: N=100 and N=200 conditions were dummy-coded and used in analyses above. *df* = 1 for all analyses above. * p<.05, ** p<.01

Table 6.

Percentage of Metric Invariance Tests Significant by Study Condition

	K=3				K=6			
	All Lower		Mixed		All Lower		Mixed	
	Lowest	Highest	Lowest	Highest	Lowest	Highest	Lowest	Highest
N100	0	1	0	7	1	0	1	0
N200	8	25	11	83	11	4	13	4
N400	74	99	89	100	64	50	71	59

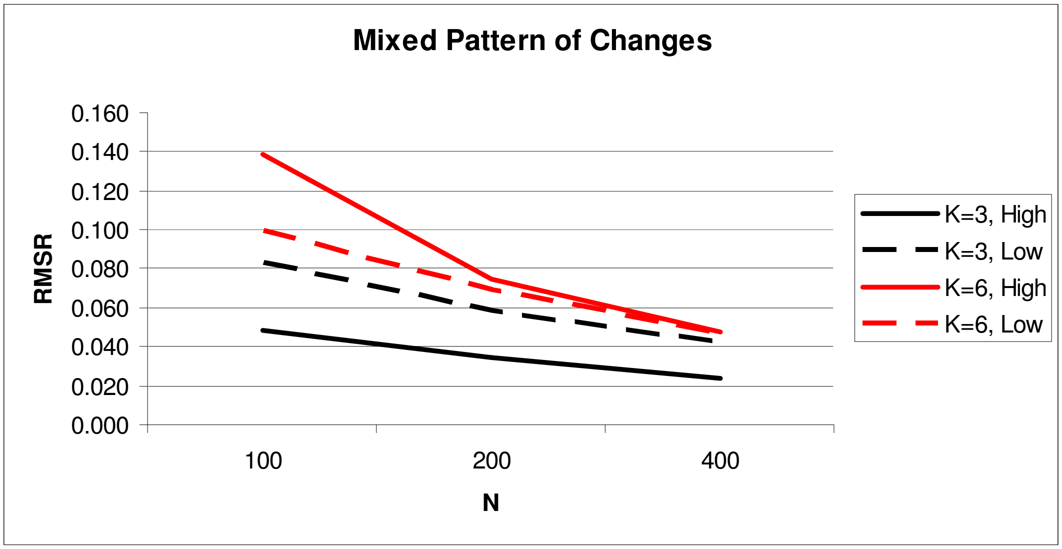
Note: K = number of factors, All lower and Mixed indicate the manner with which factor loadings were manipulated, Lowest and Highest indicate which factor loadings were manipulated.

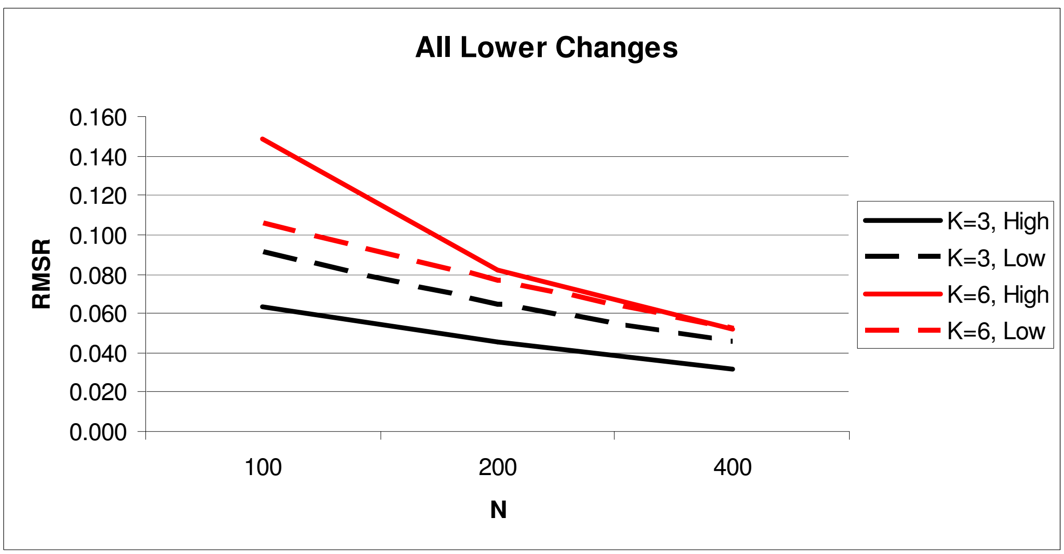
Figure 1. RMSE Values by Study Condition

Figure 2. Standard Deviation of Difference between Group 1 and Group 2 Factor Loading Estimates for Four Items with Differences

Figure 3. Percentage of Metric Invariance Tests Significant by Study Condition

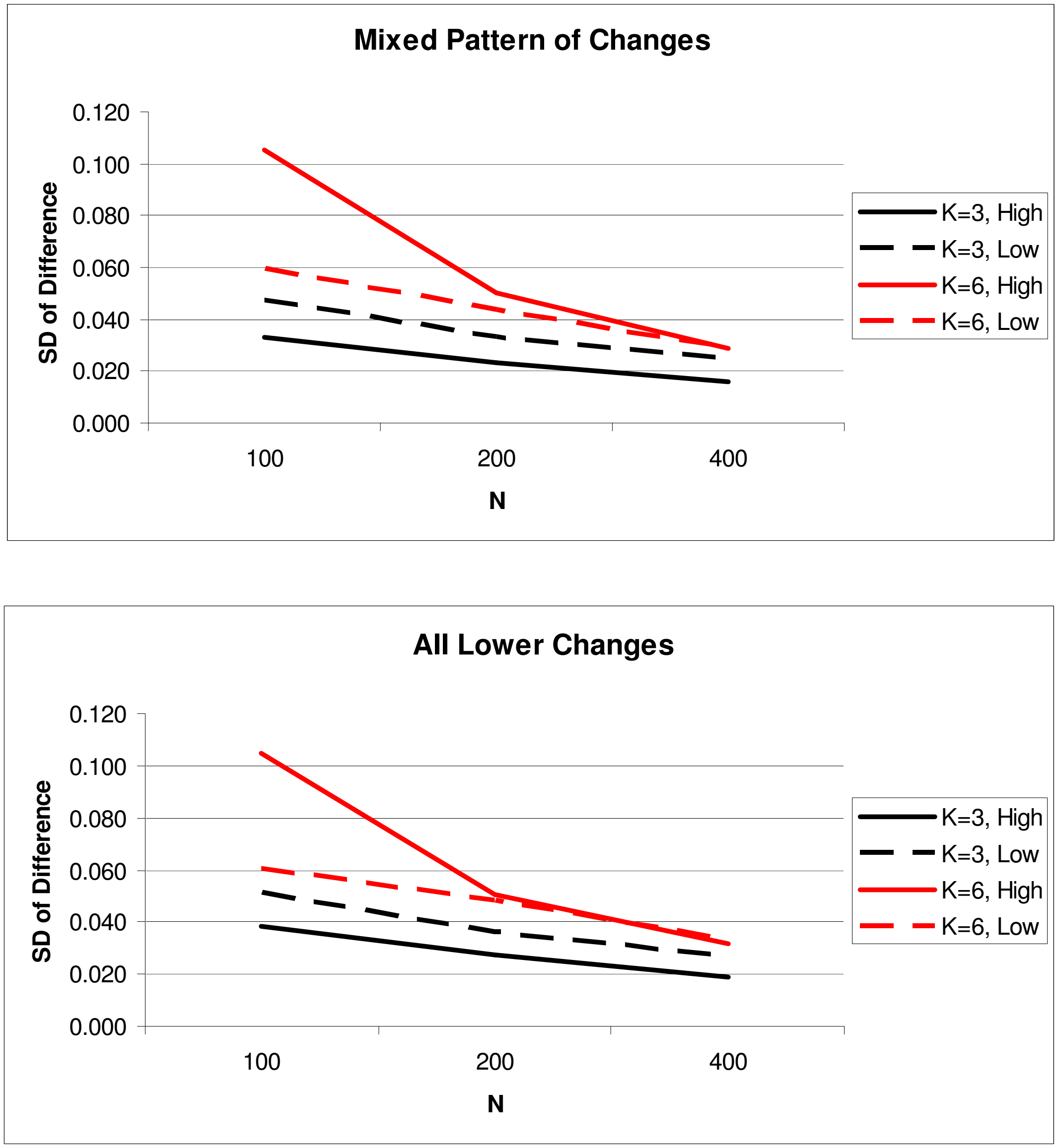
Figure 1.





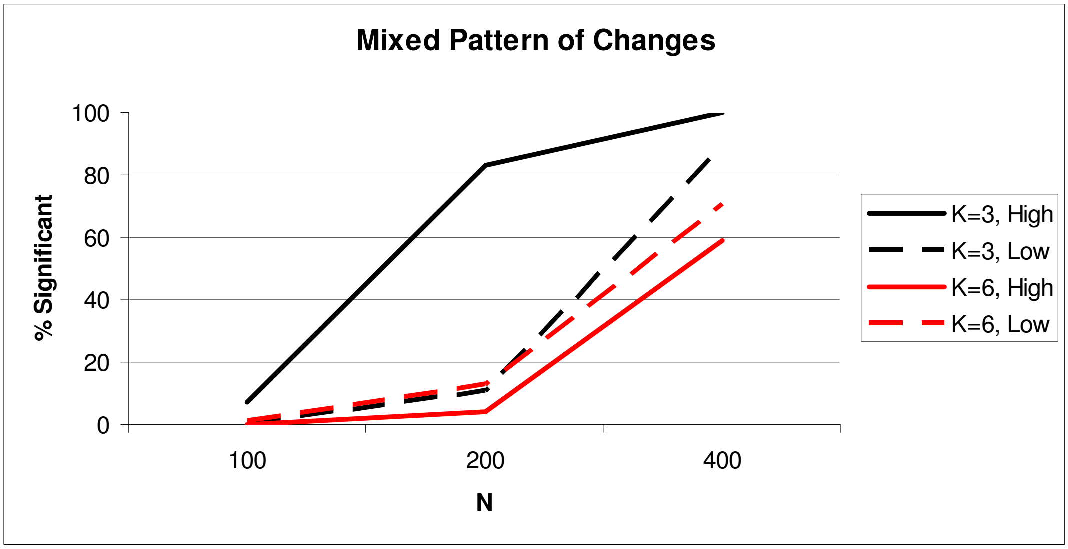
Note: K is the number of factors. High and Low reference the communality of the DF items.

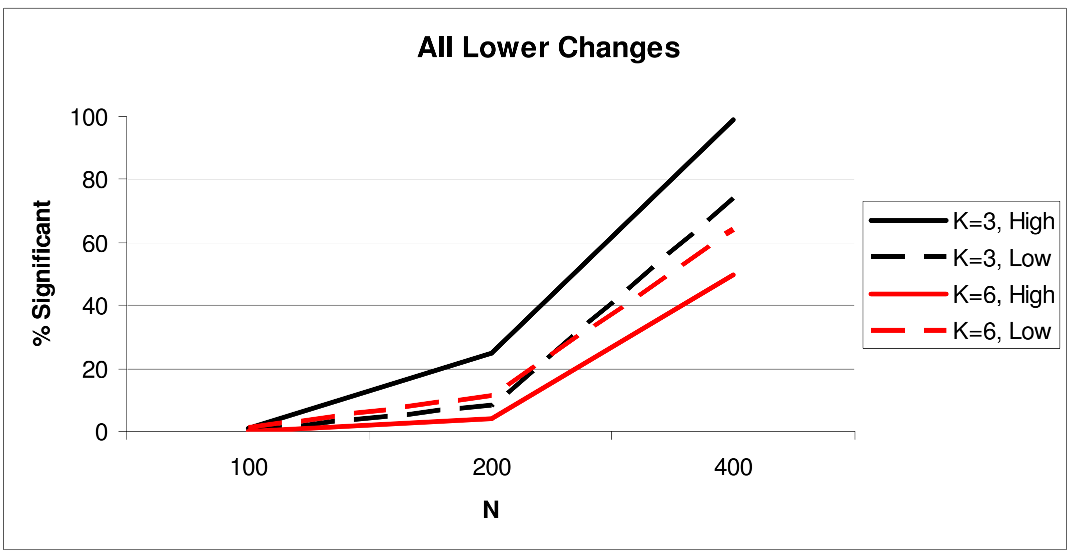
Figure 2.



Note: K is the number of factors. High and Low reference the communality of the DF items.

Figure 3.





Note: K is the number of factors. High and Low reference the communality of the DF items.