A comparison of confirmatory factor analysis methods
Stuive, Ilse

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2007

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
6.

Comparing the ability to adjust incorrect assignments: OMG versus CCF method

6.1. Introduction

If an assignment is indicated to be incorrect, either by CCF or OMG, the next question is how to obtain a better assignment. Preferably, these changes in item assignment can be justified substantively in that the items can be conceptually linked to their reassigned subtests. However, such theoretical considerations are often lacking, and the researcher has to resort to results of the statistical analyses performed. In two large scale reviews of articles that applied some variant of covariance structure analysis (Brekler, 1990; MacCallum, Roznowski & Necowitz, 1992), among which several included a CCF analysis, 37 out of 100 articles mentioned the modification of an original assignment. Among the 37, only six provided a clear theoretical basis for the modifications. The 31 remaining articles fully relied on suggestions provided by the statistical analysis. Usually, those suggestions are based on fit statistics of modified models to the sample under study.

For both the OMG and CCF methods adjustment procedures are available that provide information that can be used to adjust an assignment. The adjustment procedures for both methods will be subsequently discussed in the following sections. The goal of the studies presented in this chapter is to compare the quality of the adjustment procedures. Specifically, the main questions are: When items are assigned to subtests incorrectly, to what extent are the adjustment procedures for the OMG and CCF methods able to indicate the proper assignment, and under which conditions? This question will be answered using three simulation studies. In the first two simulation studies the adjustment procedures are compared on simulated data sets, whereas with the third study sample data sets are again drawn from the large empirical data set that was also used in the previous chapter.
6.1.1. Adjustment procedures for the OMG method

The OMG method considers for each item the observed correlations between the item and each subtest, where each subtest is an unweighted sumscore of items. An item assignment is considered to be correct when the subtest-item correlation is highest for the subtest the item is supposed to belong to and incorrect otherwise (e.g., Nunnally, 1978, p. 399). Moreover, the correlations also indicate to what subtest the items should be assigned instead: that is, it should be assigned to the subtest with which it correlates highest. As this OMG adjustment procedure consists of only one step, we will denote it as the one-step OMG procedure. We propose an alternative as well, namely an iterative adjustment procedure. In the iterative procedure the adjusted assignment obtained from an OMG analysis is tested in a subsequent OMG analysis on the same data. From the latter analysis it will either be concluded that the assignment is supported by the data, or suggestions for new adjustments are given. In the latter case, this new assignment can be tested again on the same data set. The iterative procedure may continuously switch between possible adjusted assignments when none of the adjusted assignments is supported by the data. Therefore, we stop the iterative procedure if an obtained adjusted assignment equals an adjusted assignment already obtained in one of the previous steps.

6.1.2. Adjustment procedures for the CCF method

Numerous fit indices have been proposed to study how well the model estimated by CCF fits the data. These indices may indicate that the estimated model does not fit the data well which implies that the assignment of items to subtests used is probably incorrect. Specification search strategies (Leamer, 1978) have been developed to obtain information about possible adjustments. With these specification search strategies, one or more parameters of the factor model are adjusted in such a way that an assignment of items to subtests is obtained that is better able to describe the data at hand. To decide which parameters to adjust, several tools are available, such as the multivariate Lagrange multiplier and Wald tests (Chou & Bentler, 1990), the expected parameter change statistic (Kaplan, 1989) and the modification index (Sörbom, 1989).

We focus on the modification index (MI) because it is probably most often used. The MI is also known as the univariate Lagrange multiplier test. The MI indicates for each fixed or constrained parameter the expected decrease in the overall chi-square statistic that results from the maximization of the likelihood function when
Comparing the ability to adjust incorrect assignments

this parameter would be estimated freely. A significant decrease of the overall chi-square statistic can be an indication to add this parameter to the free elements of the factor model. After finding one or more MIs with a significant chi-square, it is a common approach to add the parameter with the highest MI to the factor model and re-estimate the model using the same data set. This process continues until no significant MI is found anymore. MacCallum, Roznowski and Necowitz (1992) called this procedure the sequential model modification procedure. With this procedure it is advisable to modify one specific parameter at a time, as the parameter estimates are often not independent, implying that the addition of one parameter can influence the modification index of a second parameter. Clearly, in the sequential modification procedure the number of parameter estimates increases, resulting in less parsimonious models.

After the application of the sequential modification procedure it is important to investigate the resulting parameter estimates. Some of the estimated parameters could have become less important in explaining the data due to dependence of the parameter estimates. More parsimonious models can now be obtained by removing these relatively unimportant parameters from the model. One of the most often used procedures is to use the Wald test to decide whether or not a parameter estimate significantly differs from zero. After the sequential modification procedure, the non-significant parameters are then often simultaneously removed from the factor model to obtain a more parsimonious model (e.g., Saris & Stronkhorst, 1984; MacCallum, 1986; Green, Thompson & Poirier, 1999).

Despite the fact that specification searches are often used in practice (Breckler, 1990; MacCallum, Roznowski, & Necowitz, 1992) not much research has been done to study the quality of the search results. MacCallum (1986) and Silvia and MacCallum (1988) investigated the quality of specification searches in covariance structure analysis and reported that frequently no better assignment was obtained after a specification search. However, both studies mainly focused on the structural part of the models instead of the measurement part, whereas CCF analysis only uses the measurement part. We know of only two articles (Hutchinson, 1998; Green, Thompson, & Poirier, 1999) that evaluated the quality of search strategies using measurement models. Both studies used incorrect assignments on simulated data sets to study the quality of a specification search. From the two studies by Hutchinson and Green et al. we can distinguish two types of incorrect assignments that may occur in practice.

The first type of incorrect assignment is characterized by the ignorance of so-called secondary loadings. Hutchinson (1998) argued that in practice researchers are
often inclined to assign each item only to one subtest, whereas items could well be related to a second subtest as well. Characteristic of these so-called secondary assignments is that the item is less strongly related to this second factor compared to the other, primary, subtest. Hutchinson used simulated population data that were constructed in such a way that some of the items were assigned to two subtests instead of one. Samples were drawn from this population data and various incorrect assignments were used on the data sets. With these incorrect assignments, both primary as well as secondary assignments of items to subtest were omitted. It appeared that the specification search procedure performed relatively poorly with low sample sizes ($N \leq 400$) and when relatively many items were incorrectly not assigned to a subtest. Recovery of the assignment present in the population data increased when sample size increased. Furthermore, omitted primary assignments were detected easier by the modification indices than omitted secondary assignments. Green, Thompson and Poirier (1999) also studied the detection of secondary loadings. They found good performance of specification search procedures with this type of incorrect assignments. However, their study was limited to testing incorrect assignments on simulated population data sets. Therefore, for example, the influence of sample size on the detection of these specific secondary loadings could not be studied.

The second type of incorrect assignment is characterized by the use of an incorrect simple structure. In test construction it is desirable to obtain a questionnaire with a simple structure, where each item is solely assigned to one subtest. However, occasionally it can occur that some of these items are assigned wrongly. One of the illustrations by Green, Thompson and Poirier (1999) focused on such an incorrect assignment. They assigned three items to a subtest to which they were not assigned in the simulated population data. The procedure resulted in the addition of two correct and four incorrect parameters. So, one incorrectly assigned item was not recognized as such using the modification indices. Furthermore, none of the incorrect parameter estimates could be fixed to zero after the sequential modification procedure on the basis of non-significant Wald test outcomes, and thus the model still contained incorrectly assigned items. So, with this specific simulated data set, the specification search was not very successful. However, again this study was limited to a single population data set and one, rather extreme, type of incorrect assignment where three items were incorrectly assigned.

Unfortunately, not much is yet known about the quality of specification search procedures with incorrect simple structure assignments using the CCF method since this was only studied on one population data set. With the OMG method the first type of incorrect assignment cannot be studied as with this method always a
Comparing the ability to adjust incorrect assignments

simple structure has to be defined. For these reasons we will focus on incorrect simple structure assignments in the present study.

The goal of the current study is to compare the quality of adjustment procedures by both the OMG and CCF methods. Specifically, we compared two OMG variants, the one-step and iterative variant, with six CCF adjustment approaches, to be discussed in the Method section. We evaluated their performances under several conditions. Under certain conditions, disappointing results were obtained using the CCF method. As we suspected this could be due to the small number of items per subtest used in Study 1, we increased the number of items per subtest in a second study. In this second study, we only evaluated the CCF adjustment procedure that appeared best in Study 1, and compared it with the results of the one-step and iterative OMG adjustment procedure.

6.2. Method Study 1

6.2.1. Construction of the data

For both studies sample data were constructed for which the true assignment of items to subtests in the population was known. In Study 1, sample data were constructed from a population data set consisting of 12 items. The true assignment was specified such that the first four items were assigned to the first subtest ($s_1$), the subsequent four items to the second subtest ($s_2$) and the last four items to the third subtest ($s_3$). The data were constructed following the common factor model

$$x = \Lambda \xi + \delta.$$ 

Matrix $\Lambda$ is a simple structure loading matrix of order $12 \times 3$, containing only zeros and ones. The ones in $\Lambda$ indicate the true assignment of items to subtests present in the population. The factor scores in $\xi$ were drawn at random from a multivariate normal distribution $N(0, \Phi)$, where $\Phi$ is a pre-specified correlation matrix of the three common factors. All correlations between the common factors were fixed at 0.3. The random unique scores were drawn from a multivariate normal distribution $N(0, \Theta)$, where matrix $\Theta$ is a diagonal matrix of order $12 \times 12$ with on the diagonal the unique variances $\sigma^2$ of the 12 items.

The unique variances and sample sizes were varied, namely the unique variances were varied such that 20%, 50% or 80% of the total amount of variance explained by the items was accounted for by the unique part $\delta$ and the sample sizes were taken equal to 100, 200, 400 or 1000. The design was completely crossed and replicated 100 times so that the total number of data sets analyzed was 3 (unique
variance percentages) × 4 (sample sizes) × 100 (replications) = 1200 data sets, which were generated using MATLAB7 (MathWorks, 2004).

For each simulated data set we considered a mild and a strong incorrect simple structure assignment that were both based on the true assignment. The two types are:

1. moving 1 item from \(s_1\) to \(s_2\) (mild)
2. moving 1 item from \(s_1\) to \(s_2\) and 1 item from \(s_2\) to \(s_1\) (strong)

In the simulation study correlation matrices were analyzed. The one-step and iterative OMG methods were performed using our own script file, running under MATLAB7. For the CCF method, we used the LISREL program (Jöreskog & Sörbom, 2001) using the Maximum Likelihood procedure. For the CCF method we selected six adjustment procedures as will be discussed now.

### 6.2.2. Six adjustment procedures for the CCF method

For our simulation study, we selected those CCF adjustment procedures that are most often used in practice. MacCallum, Roznowski and Necowitz (1992) noted that, in practice, researchers often use a rather mechanical approach to modify their models, sequentially adding the parameter with the highest MI, until no significant MI is found or until a good model fit is obtained. With more than one significant MI, parameters with the highest MI are often added first. An adjustment procedure which closely matches this approach is the automatic modification option available within the LISREL program, which was also used in the simulation studies by MacCallum (1986), Silvia and MacCallum (1988) and Hutchinson (1998). All six CCF procedures used in our simulation study were variants of this automatic modification procedure. The six CCF procedures differ in which parameters are allowed to be set free (\(\Lambda\) and/ or \(\Theta\)), and in the use of a stopping criterion. An overview of the CCF adjustment procedures and the names as adopted here is given in Table 6.1. First, we will describe two different search strategies that differ in the parameters that are allowed to be set free. Next, three different stopping criteria will be discussed.

Without further specifications, the LISREL automatic modification option is allowed to add parameters in \(\Lambda\) and \(\Theta\). However, the off-diagonal elements in \(\Theta\) correspond with the correlations between the unique parts of the items, which in ordinary CFA models are assumed to be zero, unless theoretical grounds indicate otherwise. Moreover, Fornell (1983) and Bagozzi (1983) have noted that in practice
Table 6.1. CCF adjustment procedures used with the automatic modification procedure.

<table>
<thead>
<tr>
<th>Stopping criterion</th>
<th>Search strategy</th>
<th>Restricted</th>
<th>Unrestricted</th>
</tr>
</thead>
<tbody>
<tr>
<td>No stop</td>
<td>( \Lambda ), No stop</td>
<td>( \Lambda ), ( \Theta ), No stop</td>
<td></td>
</tr>
<tr>
<td>RMSEA ( \leq .03 )</td>
<td>( \Lambda ), RMSEA</td>
<td>( \Lambda ), ( \Theta ), RMSEA</td>
<td></td>
</tr>
<tr>
<td>RMSEA ( \leq .05 ) or SRMR ( \leq .06 )</td>
<td>( \Lambda ), Comb</td>
<td>( \Lambda ), ( \Theta ), Comb</td>
<td></td>
</tr>
</tbody>
</table>

the correlations between the unique parts are often added to a factor model without providing a theoretical justification. These modifications are then solely made to improve the fit of the factor model. A lack of theoretical justifications for making modifications was also reported by Breckler (1990) and MacCallum, Roznowski and Necowitz (1992). Silvia and MacCallum (1988) indicated that when modifications were restricted to those that do not violate the structure of the population model, the quality of the adjustment procedure is improved. Despite these arguments against allowing the addition of off diagonal elements in \( \Theta \), it is still used in the automatic modification procedure in practice. Therefore, we decided to compare the restricted adjustment procedure, where only modifications of fixed elements in \( \Lambda \) were allowed (and not in \( \Theta \)), with the unrestricted adjustment procedure allowing modifications in \( \Lambda \) and \( \Theta \).

Default, the automatic modification continues adding fixed parameters corresponding to significant MIs to the factor model until no significant (using the .05 level) MI is found. This default procedure does not stop prematurely when a good model fit is obtained and will therefore be denoted as the “No stop” condition.

The previous studies by MacCallum (1986), Silvia and MacCallum (1988) and Hutchinson (1998) included the use of a stopping criterion. In practice, researchers might stop an automatic modification prematurely when a good model fit is already obtained in the previous modification step. Hutchinson, MacCallum and Silvia and MacCallum decided to stop the automatic modification procedure prematurely when a non-significant chi-square value was obtained. However, they all indicated that, on some occasions, this resulted in a premature stop of the adjustment procedure where additional modifications were still necessary to obtain the true assignment. Mainly because of these findings, but also because of its sensitivity to sample size, we decided not to use the chi-square stopping criterion. Instead, we chose an index with a specific criterion value of which previous research has indicated that it results in a good detection of incorrect assignments. In Chapter 5 we
found relatively good results when using a combination of the RMSEA>.05 and SRMR>.06 to indicate an incorrect assignment. Also good results were obtained when the RMSEA>.03 was used. Therefore, as two alternative adjustment procedures, we used the LISREL automatic modification option in conjunction with RMSEA≤.05 and SRMR≤.06, and RMSEA≤.03, respectively, as stopping criteria. For the use of the combination of RMSEA and SRMR this means that the modification process continues until for at least one fit index a value lower than or equal to, respectively, .05 or .06 is obtained.

6.2.3. Defining the dependent variable

For each of the two incorrect simple structure assignments we studied the frequency with which the six CCF adjustment procedures and both OMG adjustment procedures were able to correctly indicate how to adjust the incorrect assignments in such a way that the true assignment was obtained. Then, for both the CCF and OMG method, proportions correct were computed for each cell in the design.

With the OMG method, an obtained simple structure was considered correct when it equaled the simple structure of the true assignment and incorrect otherwise. For the iterative OMG method we also wanted to see how many iterations were needed for the method to obtain a stable solution. It can be derived that a minimum of three iterations is needed for the method to indicate the true assignment. With the first iteration the incorrect simple structure is used on the data. This will ideally result in obtaining an adjusted assignment that is equal to the true assignment which is subsequently used on the data in the second iteration. As defined previously, the iterative procedure stops when an adjusted assignment is equal to one of the previous assignments. Therefore, the procedure will stop in this case after using the true assignment for the second time on the data, which occurs at the third iteration. In the Results section we will show how often three iterations are indeed sufficient to obtain the true assignment and under which conditions more iterations are needed. We also show how often these additional iterations are successful in obtaining the true assignment.

For the CCF method several criteria for assessing correct recovery are possible. Of these criteria, we decided to choose the most liberal one, namely to consider a CCF judgment as correct when the modifications suggested by the automatic modification procedure led to a model assigning all items to at least the factor it should be assigned to. For example, our first incorrect simple structure
Comparing the ability to adjust incorrect assignments

assignment type was created by assigning one item from the first to the second factor. We considered a judgment as correct when the suggested adjustments provided by the automatic modification procedure included the suggestion to assign the incorrectly assigned item to the first subtest. Thus, we ignored the presence of the incorrectly estimated parameter between that item and the second subtest, and also the possible presence of additional, incorrectly suggested, parameter estimates in \( \Lambda \).

6.3. Results Study 1

Each simulated data set was analyzed using the two incorrect simple structure assignment types. The resulting 1200 × 2 = 2400 analyses were performed using each of the six CCF adjustment procedures and the two OMG adjustment procedures, which results in a total number of 2400 × 8 = 19200 analyses. However, during the analysis using the CCF method, occasionally the LISREL program did not converge or gave another type of warning. When this occurred, no modification indices were provided. In total, 1851 of the 14400 CCF analyses resulted in a warning. These warnings occurred roughly equally frequently across the two search strategies, namely 921 warnings with the restricted search and 930 for the unrestricted search. Within the two search strategies, the frequency of warnings somewhat increased from the “Comb” condition to the “No stop” condition. This makes sense as the chance of obtaining a warning increases with the continuation of the automatic modification procedure. When we studied a small selection of the cases that gave these warnings we found that, surprisingly, most of these warnings were due to the automatic modification option. When we sequentially modified the models manually, no non-convergence warnings or other warnings occurred for this small selection. We concluded that these warnings were possibly due to an implementation error of the automatic modification option and therefore chose to discard all specific data sets for which warnings occurred from further analysis. To ensure that the results of the OMG method were based on a similar selection of data sets, we decided also to remove corresponding data sets for the OMG method. In total 317 data sets were removed for both methods, but we want to note here that the OMG method performed rather well on these data sets. Both the one-step and iterative OMG adjustment procedures performed reasonably well, with a perfect recovery of the true assignment for 242 and 274 of the 317 data sets, respectively.

The proportions correct for the two incorrect assignment types and their 95% confidence intervals are plotted in Figure 6.1 for the six CCF and two OMG
adjustment procedures. From Figure 6.1 we see that higher proportions correct were obtained consistently with the mild incorrect assignment. In that condition, the two OMG adjustment procedures and the two CCF adjustment procedures where no stopping criterion was used performed best. With the strong incorrect assignment the iterative OMG adjustment procedure outperformed the one-step OMG adjustment procedure as well as all six CCF adjustment procedures.

**Figure 6.1.** Averaged proportions correct for the six CCF and two OMG adjustment procedures when they were used with the two incorrect assignment types.

We decided to average results across the two incorrect assignment conditions and then to present the results in relation to the two other variables varied in this study: sample size and the amount of unique variance. Figures 6.2a, b and c give, for each proportion of unique variance, 20%, 50% and 80% respectively, the proportions correct for the six CCF adjustment procedures and the two OMG adjustment procedures as a function of sample size.

From Figure 6.2a we see that, in the 20% unique variance condition, both OMG adjustment procedures very often yielded the correct assignment for all four sample sizes. For the CCF method we see, not surprisingly, an increase in the proportion correct when sample size increased. We also see that restricting the
Figure 6.2. Proportions correct for the six CCF methods and the two OMG methods as a function of sample size in the case of 20% (Figure 6.2a), 50% (Figure 6.2b) or 80% (Figure 6.2c) unique variance.
adjustment procedure to elements in \( \Lambda \) (first three bars) resulted in higher proportions correct than allowing modifications of elements in \( \Lambda \) as well as in \( \Theta \) (fourth to sixth bars). The differences between the two search strategies somewhat decreased with increasing sample size. Virtually no differences in proportions correct were found between the three stopping criteria when the percentage of unique variance was fixed at 20%.

With 50% unique variance, both OMG adjustment procedures performed almost perfectly, as illustrated in Figure 6.2b. Only when samples of size 100 were drawn, proportions correct lower than one were obtained with the one-step OMG adjustment procedure. For the CCF method, we again see higher proportions correct with the procedure not allowing modifications in \( \Theta \). In contrast with Figure 6.2a, we see differences between the effects of the three stopping criteria when samples of size 100 were analyzed. Lower proportions correct were obtained when the combination of \( \text{RMSEA} \leq 0.05 \) or \( \text{SRMR} \leq 0.06 \) was used compared to the other two stopping criteria.

When the percentage of unique variance was increased to 80%, we see larger differences in proportions correct between the different procedures, as shown in Figure 6.2c. With the iterative OMG adjustment procedure, consistently higher proportions correct were obtained than with the one-step OMG adjustment procedure. Proportions correct near one were obtained with the OMG adjustment procedures only when sample sizes were at least 400 (for iterative OMG) or 1000 (for one-step OMG). The CCF method only resulted in proportions correct near one when no stopping criterion was used and sample size was 1000. Even then, this CCF procedure performed worse than the OMG adjustment procedures. The restricted adjustment procedure again performed better than the unrestricted adjustment procedure, though differences were much smaller than with 20 or 50% unique variance. In contrast, differences due to the three stopping criteria increased when going from 20 to 80% unique variance. Much higher proportions correct were obtained when no stop was used compared to the other two stopping criteria.

For the iterative OMG adjustment procedure we also studied the number of iterations that were needed to obtain a stable solution. We found that with 20% and 50% unique variances three iterations were always sufficient to obtain a stable iterative OMG solution. In Section 6.2.3 we already indicated that this was the minimum number of iterations needed to adjust the incorrect assignment in such a way that the true assignment was obtained. With 80% unique variance, the number of iterations necessary for a stable solution could vary from 3 up to 12 iterations. The highest number of iterations was observed in this condition with samples of size 100. With higher sample sizes three or four iterations were often sufficient for the iterative
Comparing the ability to adjust incorrect assignments

procedure to stop. As for the performance of the iterative OMG method using more than three iterations, we found that with an increase in sample size the true assignment was obtained more often. With sample sizes of 400 and higher the true assignment was indicated always.

Overall, the iterative OMG adjustment procedure performed better than the one-step OMG adjustment procedure. Furthermore, of the six CCF adjustment procedures, only allowing modifications in $A$ and no stopping criterion seemed the best procedure. However, both the iterative OMG adjustment procedure and the best CCF adjustment procedure resulted in low proportions correct with 80% unique variance and sample sizes of 200 and lower. It could be the case that these lower proportions of correct judgments were due to the small number of items per subtest. Therefore, we decided to perform a second simulation study where six instead of four items per subtest were specified. To keep the study clear, we compared the best CCF adjustment procedure from Study 1, “A, No stop”, with the two OMG adjustment procedures.

6.4. Method Study 2

To study whether the low proportions correct found in Study 1 were due to the small number of items per subtest used in that study, we performed a second simulation study where six instead of four items per subtest were specified. Specifically, in Study 2 we used exactly the same design as in Study 1, except for the loading matrix. The loading matrix was specified such that, with the true assignment, the first six items were assigned to the first subtest ($s_1$), the subsequent six items to the second subtest ($s_2$) and the last six items to the third subtest ($s_3$).

Additionally, a third incorrect simple structure assignment was used on a selection of the data sets to investigate the performance of the three methods using a more serious incorrect assignment (to be noted as ‘heaviest’). On the data sets constructed with 50% unique variance, two items of $s_1$ were assigned to $s_2$ and one item of $s_2$ was assigned to $s_1$. We chose the 50% unique variance condition because it appeared to be a condition with ‘medium’ performance. Sample sizes were varied as in Study 1 and 100 replications were used. The third incorrect simple structure assignment was therefore used on 100 (replications) $\times$ 4 (sample size) = 400 data sets with items containing 50% unique variance. Therefore, in $2400 + 400 = 2800$ analysis were performed by the three adjustment procedures.
6.5. Results Study 2

As in Study 1, the LISREL program reported warnings rather often in the performance of a CCF analysis. As in this case no modification indices were provided these data sets needed to be discarded. Out of the 2800 analyses, 125 resulted in a warning. All were removed from further analysis for both the CCF and OMG methods. The use of the one-step OMG adjustment procedure resulted in a perfect recovery of the true assignment on 90 of these 125 data sets whereas the iterative OMG adjustment procedure was able to recover the true assignment in 101 of the 125 data sets.

We present the results after averaging over the first two types of incorrect simple structure assignments as in Study 1, so as to make results optimally comparable to those presented in Figure 6.2. However, because proportions correct were similar for the three methods in the conditions with 20% and 50% unique variance, we only show the results for the 50% and 80% unique variance conditions. Figures 6.3a and b give for both unique variance conditions respectively, the proportion correct for the three adjustment procedures using the four different sample sizes. Figure 6.4 shows the proportions correct for each adjustment procedure when the heaviest incorrect simple structure assignment was used on data sets consisting of items with 50% unique variance.

We see from Figures 6.3a and b that the proportions correct by the three adjustment procedures only deviate from one when the amount of unique variance was fixed at 80%. From Figure 6.3b we see that proportions correct are only equal to one in this unique variance condition when samples of size 1000 were analyzed and nearly one with samples of size 400. We also see that, especially for the lower sample sizes, the proportions correct obtained by the iterative OMG adjustment procedure were consistently higher than those obtained by the one-step OMG and CCF procedure.

As for the number of iterations used with the iterative OMG method it appeared that the minimum of three iterations was exceeded less often compared to Study 1. Only when the amount of unique variance was equal to 80% in combination with sample sizes of 200 and lower occasionally more than three iterations were needed. Often these additional iterations were necessary for the iterative OMG method to obtain the true assignment.
Comparing the ability to adjust incorrect assignments

Figures 6.3a and b. Proportions correct for the CCF adjustment procedure and the two OMG adjustment procedures as a function of sample size for 50% (Figure 6.3a) and 80% (Figure 6.3b) unique variance when the mild and strong incorrect assignments were used.
When we study the results using the third assignment condition summarized in Figure 6.4, we see that again proportions correct equal to one were obtained with all three methods when samples of sizes 400 and 1000 were used. With lower sample sizes, the use of the iterative OMG adjustment procedure still resulted in proportions correct equal to one whereas the use of one-step OMG procedure only resulted in the perfect recovery of the true assignment when samples of size 200 or higher were used. The use of the CCF method resulted in lower proportions correct for sample sizes of 200 and lower. All in all, this study shows that results get better when the number of items is increased, while the pattern of differences between methods and conditions remains similar.

6.6. Method Study 3

In the third study we used the same data set as used in Section 5.5. The total data set consisting of 2860 observations was again considered to be the population data set. From this data set samples of sizes 100, 200, 400 and 1000 were drawn. For each sample size 10 replications were created which results in 4×10 = 40 different data sets. On these data sets the five incorrect assignments given in Table 5.5 were used.
These 40×5=200 analyses were performed using again the two OMG and six CCF methods also used in Study 1 of this chapter.

6.7. Results Study 3

Data sets leading to warnings in CCF were removed from further analysis. In total 28 data sets were removed of which 16 were judged correctly using the iterative OMG method and 15 using the one-step OMG method.

Figure 6.5 shows for each sample size the proportions correct (with 95% confidence intervals) for the two OMG and six CCF methods. Compared to previous studies, the confidence intervals are rather broad. This is mainly because of the smaller number of replicates. From Figure 6.5 we see that the iterative OMG method always results in higher proportions correct compared to the one-step OMG method or the six CCF methods. Also high proportions correct were obtained using the restricted search procedure for the CCF method. With higher sample sizes, differences between the stopping criteria increased, where “No stop” and RMSEA≤.03 resulted in higher proportions correct compared to the combination of RMSEA≤.05 or SRMR≤.06. These results are rather similar to those presented earlier in this chapter.

![Figure 6.5. Proportions correct for the six CCF and two OMG methods as a function of sample size when using the empirical data set.](image-url)
6.8. Discussion

This study compared six CCF adjustment procedures with two OMG adjustment procedures on their ability to adjust incorrect simple structure assignments. Here, we will first discuss the six CCF adjustment procedures after which we will compare the best of these six CCF adjustment procedures with the two OMG adjustment procedures.

From the comparative results of Study 1, for the CCF method, we conclude that restricting the adjustment procedure to elements in $\Lambda$ results in higher proportions correct than allowing modifications of elements in $\Lambda$ as well as in $\Theta$. This finding confirms results of previous studies by Silvia and MacCallum (1988) and MacCallum (1986). They already indicated superior results for the restricted adjustment procedure. Also, Gerbing and Anderson (1984) indicated that the addition of correlations between errors could mask the known underlying structure. They also noted that “while use of correlated measurement errors improves fit …, it does so at a correspondent loss of the meaning and substantive conclusions which can be drawn from the model” (p. 574). Without having a theoretical basis for such correlations it seems therefore preferable to limit the adjustment procedure to elements of $\Lambda$.

We also found that a fit-measure based stop of the adjustment procedure, based on good model fit, resulted in lower proportions correct. MacCallum (1986), Silvia and MacCallum (1988) and Hutchinson (1998) already mentioned that the use of a non-significant chi-square value resulted in a premature stop where additional modifications were still necessary to obtain the true assignment. We made the same observation when using values of the RMSEA and SRMR indices as stopping criteria. Therefore, we advise not to use a premature stop or, if necessary, to choose a rather liberal stopping criterion.

Comparing the best CCF adjustment procedure with the two OMG adjustment procedures we found that the iterative OMG adjustment procedure worked better or comparable to the CCF procedure. Especially with higher amounts of unique variance and more difficult incorrect assignments the iterative OMG method frequently outperformed the CCF method. Not surprisingly, this difference decreased with increasing sample size. Results of the one-step OMG procedure were also rather comparable to those obtained with the CCF method. Only with 80% unique variance the CCF method clearly outperformed the one-step OMG procedure. For both methods overall slightly higher proportions correct were obtained when the number of items per subtest increased.
Thus, these studies gave rather similar results for the iterative OMG method and the best CCF approach. However, we should emphasize that we used a rather liberal approach for the CCF method. In this study we only focused on minimizing the chance of making a Type II error (failure to identify a parameter that should be included in the CCF model). We ignored the fact that several parameter estimates in \( \Lambda \) were added to the model by the automatic modification procedure whereas they were zero in the true assignment. Furthermore, the parameter estimates corresponding to the incorrect assignment were not removed. So, with the CCF method the simple structure of the true assignment was not by definition obtained with a correct judgment and proportions correct could be considerably lower when a less liberal approach would be used.

There are several procedures available to remove incorrectly added parameter estimates, which are also known as Type I errors. We tested a rather simple procedure, namely we studied the proportions correct when for each item the highest absolute parameter estimate in \( \Lambda \) was chosen. The obtained simple structure was then compared with the true assignment. Only slightly lower proportions correct were obtained with this procedure for data sets with 80% unique variance and low sample sizes. Therefore, this could be a good strategy to remove incorrect parameters and obtain a more parsimonious model.

Another alternative to remove the incorrectly added parameters is to use the Wald test. With this test nonsignificant parameter estimates can be removed from the CCF model. Usually, these parameter estimates are removed when the Wald test indicates significance above the .05 level. Green, Thompson and Poirier (2001) indicated that under certain conditions a better control for Type I error is obtained with the Wald test using an adjusted Bonferroni correction method. Furthermore, Green, Thompson and Babyak (1998) indicated that Type I errors also could be remedied to some extent using a Bonferroni correction during, instead of after, the automatic modification procedure. Their study showed that the Bonferroni correction reduced the number of incorrectly added parameter estimates. Unfortunately, this was accompanied by an increase of Type II errors. Further research is necessary to see which procedure provides the best balance between both types of errors.

It should also be noted that in this study data were generated such that continuous item scores were obtained. In practice the item scores are often categorical in nature. Both methods will probably work less well when the number of categories is rather small. It can be expected that under these conditions the OMG method outperforms the CCF method as with this method no specific assumptions about the item distributions are made. However, it can also be argued that with a small number
of item categories an alternative CCF analysis should be used that is specifically
designed to deal with this type of data. This method, known as the asymptotically
distribution free (ADF; Browne, 1984) method, can be used to analyze categorical
data but only when relatively large sample sizes are available. Further research is
necessary to see how well the methods perform with a small number of item
categories.

Furthermore, data were constructed using equal weights for all items. This is
in accordance with the assumption of equal weights with the OMG method. It can be
the case that the data construction procedure has favored the OMG method. However,
in a small simulation study where we varied the amounts of unique variance across
the items we observed no clear difference with the results discussed here. In general
the OMG method appeared to perform well with low (20%) to moderate (50%)
amounts of unique variance. Performance decreased when items with 80% unique
variance were included. So, also with different item weights similar results were
obtained with the OMG method.

Overall, of the procedures compared in this study the iterative OMG
adjustment procedure performs better or highly comparable to the best CCF
adjustment procedure and is therefore a good alternative to the widely used CCF
method for the detection and adjustment of incorrect assignments.