

A MODEL SELECTION APPROACH TO TESTING DEPENDENT ICCs

Wolfgang Viechtbauer and David Budescu

ABSTRACT

In this commentary, we discuss the tutorial by Cohen and Doveh (this volume) on using multi-level models to estimate and test hypotheses about (dependent) intra-class correlations (ICCs). The goal in this context is to find subsets of homogeneous ICCs across time points and/or experimental tasks. We suggest that one should approach this model selection problem by (a) specifying the set of all possible models, (b) using a systematic top-down selection strategy that avoids possible inconsistent and intransitive patterns in the results, and (c) using the entire dataset available when fitting each model. Our goal is to discuss the larger framework in which the specific models considered by Cohen and Doveh in their analysis are embedded, which should be useful in guiding researchers faced with a similar analysis task.

INTRODUCTION

Cohen and Doveh (this volume) provide a useful tutorial for using multi-level models to estimate and test hypotheses about (dependent) intra-class correlations (ICCs). Specifically, the authors show how to specify a model

Multi-Level Issues in Strategy and Methods
Research in Multi-Level Issues, Volume 4, 433–454
Copyright © 2005 by Elsevier Ltd.
All rights of reproduction in any form reserved
ISSN: 1475-9144/doi:10.1016/S1475-9144(05)04019-1

that allows the ICC to differ over time and/or experimental tasks and then illustrate how hypotheses about changes in the ICC (i.e., in group homogeneity) can be examined by testing the equivalence of certain variance components in the model. Many readers can benefit from their straightforward and easy to read introductions, the (simulated and real) examples, and their careful and detailed description of the computer programs (and their outputs) used in the analyses.

In the present commentary, we will focus mostly on the analysis of the “motivating example.” In this study, participants were pre- and post-tested under two different experimental tasks, leading to four different experimental conditions. Therefore, if the ICC depends on the experimental condition, then four possible ICCs are at most possible (one for each condition), implying that the ICC depends not only on the experimental task, but also on time. This most general model, however, might be unnecessarily complex. By constraining subsets (or all) of the four ICCs equal to each other, we obtain 14 possible subset models, each of which might provide a more parsimonious description of the data.

The problem then is the selection of one of these models. We illustrate how each of these models can be fit using the SAS MIXED procedure and propose a systematic model selection strategy that differs to some extent from the one taken by Cohen and Doherty (this volume). First of all, we choose a top-down approach, starting with the most complex model and then examine the feasibility of reductions in model complexity. We recommend a top-down approach, because we consider it more prudent to err on the side of selecting a model that is too complex, instead of one that is too simple. Second, we advocate using the full dataset available when fitting each model, as opposed to using only that part of the data that is just sufficient for establishing the homogeneity of various ICCs. Third, by outlining (and fitting) all possible models, one can easily see how the various models are nested within each other. Accordingly, one can choose a sequence of model comparisons that avoids inconsistent and intransitive patterns (such as finding that the first and second and the second and third ICC are homogeneous, but not the first and third). Finally, given the computational ease with which mixed models can nowadays be fitted, it is quite feasible to simply fit all possible models and select based on, for example, Akaike’s information criterion (AIC) the one that strikes an optimal balance between model fit and complexity (this approach would be consistent with the $\min(AIC)$ criterion discussed by Dayton, 2003; Bozdogan, 1987).

However, whether one approaches the model selection problem through a series of deviance (i.e., likelihood ratio) tests involving properly nested

models or employs an information criterion such as the AIC is less critical. More importantly, one should clearly define the set of possible models under consideration, which guides either procedure and facilitates the interpretation of the results. In essence, our purpose in writing this commentary is not to critique the specific approach chosen by Cohen and Doveh (in fact, their conclusions for the motivating example coincide with those obtained by us), but instead to provide some suggestions and guidelines for researchers faced with a similar analysis task.

To be more precise in discussing the various models, we will use the notation of the general linear mixed-effects model (GLMM) in our commentary. We therefore define the GLMM in the next section and then apply its notation to the one-way random-effects model and the two-factor model with repeated measures on one factor for illustration purposes. Next, the motivating example is introduced in the same manner and the various subset models are defined. The selection of one of these models is then a straightforward matter of applying either a series of deviance tests or employing the $\min(AIC)$ criterion. However, before settling on one model, we first consider a set of more general models in which the residual variances may also depend on the condition level. In the same manner as before, we again define and fit all possible models and then select a final model. However, model selection with heterogeneous residual variances involves some additional complexities (in particular, for testing the homogeneity of ICCs) which are discussed in the next to last section. Some general comments conclude our commentary.

THE GENERAL LINEAR MIXED-EFFECTS MODEL

To discuss the various models presented by Cohen and Doveh (this volume) and in our commentary, it is helpful to write out in explicit detail the models and the implied covariance structures using the notation of the GLMM. The general form of GLMMs (e.g., Searle, Casella, & McCulloch, 1992) is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{e}$$

where \mathbf{y} is the vector of observations, \mathbf{X} the design matrix for the fixed effects parameters contained in $\boldsymbol{\beta}$, \mathbf{Z} the design matrix for the random effects contained in $\boldsymbol{\gamma}$, and \mathbf{e} the vector of random residuals. We assume $E[\mathbf{y}] = \mathbf{0}$, $E[\mathbf{e}] = \mathbf{0}$, and $\text{Cov}[\mathbf{y}, \mathbf{e}] = \mathbf{0}$. Define \mathbf{D} as the covariance matrix of the random effects parameters in $\boldsymbol{\gamma}$ and \mathbf{R} as the covariance matrix of \mathbf{e} . Then

$$\text{Var} \begin{bmatrix} \gamma \\ e \end{bmatrix} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

and \mathbf{V} , the covariance matrix of \mathbf{y} , is equal to $\mathbf{V} = \mathbf{ZDZ}' + \mathbf{R}$.

One-Way Random-Effects Model

As a simple example, consider first the one-way random-effects model that is discussed by Cohen and Doveh (this volume) and given by the authors in Eq. (1). For $i = 1, \dots, n$ groups with $j = 1, \dots, k_i$ members, the model can be written as

$$Y_{ij} = \mu + r_i + w_{ij}$$

where $r_i \sim N(0, \sigma_r^2)$, $w_{ij} \sim N(0, \sigma_w^2)$, and all the terms are independent. Then for $n = 2$ and $k_1 = k_2 = 2$ (the “smallest” case that still illustrates the important features of this model), we can write the model in GLMM notation as shown below:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [\mu] + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} + \begin{bmatrix} w_{11} \\ w_{12} \\ w_{21} \\ w_{22} \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{e}$$

Then

$$\mathbf{D} = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_r^2 \end{bmatrix}$$

and

$$\mathbf{R} = \begin{bmatrix} \sigma_w^2 & 0 & 0 & 0 \\ 0 & \sigma_w^2 & 0 & 0 \\ 0 & 0 & \sigma_w^2 & 0 \\ 0 & 0 & 0 & \sigma_w^2 \end{bmatrix}$$

Therefore, \mathbf{V} , the implied variance-covariance matrix for this model, is given by

$$V = \left[\begin{array}{cc|cc} \sigma_r^2 + \sigma_w^2 & \sigma_r^2 & 0 & 0 \\ \sigma_r^2 & \sigma_r^2 + \sigma_w^2 & 0 & 0 \\ \hline 0 & 0 & \sigma_r^2 + \sigma_w^2 & \sigma_r^2 \\ 0 & 0 & \sigma_r^2 & \sigma_r^2 + \sigma_w^2 \end{array} \right]$$

We note, as expected, that observations from different groups are uncorrelated (since the upper right and lower left (2×2) blocks of V consist entirely of zeros), while the correlation of observations from the same group is given by

$$ICC = \frac{\sigma_r^2}{\sigma_r^2 + \sigma_w^2}$$

Two-Factor Model with Repeated Measures on One Factor

Next, consider the model used in examples 4 and 5. The model, as given by Cohen and Doherty (this volume) in Eq. (8), is repeated below:

$$Y_{i(j)t} = \mu + (time)_t + (grp)_i + (id)_{j(i)} + (grp \times time)_{it} + e_{i(j)t}$$

where $Y_{i(j)t}$ denotes the t th observation from subject j in group i . This model is commonly used to analyze data from two-factor designs with repeated measures on one factor. Discussions of this model can be found, for example, in Kirk (1995), Maxwell and Delaney (2004), and Neter, Kutner, Nachtsheim, and Wasserman (1996). This type of experimental design also falls under the general category of split-plot designs.

Time (the repeated or within-subjects factor) is considered a fixed effect in this model. The group effect (the between-subjects factor) and the subject effect are random with variances σ_g^2 and σ_s^2 , respectively. The group by time interaction is also random and the variance of this effect can depend on the time point. Consequently, if there are T different time points, then in the most general case, we would need T different parameters, namely $\sigma_{gt_1}^2, \dots, \sigma_{gt_T}^2$, to describe the variances of the interaction effect. Finally, let σ_e^2 denote the variance of the residuals.

Now assume we observe $j = 1, 2$ subjects in $i = 1, 2$ groups at $t = 1, 2$ time points or measurement occasions. This is again the “smallest” case that still illustrates the important features of this model. Writing the model in GLMM notation yields:

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \frac{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}{\begin{bmatrix} \mu \\ t_1 \\ t_2 \end{bmatrix}}$$

$$+ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} g_1 \\ g_2 \\ s_{11} \\ s_{21} \\ s_{12} \\ s_{22} \\ gt_{11} \\ gt_{12} \\ gt_{21} \\ gt_{22} \end{bmatrix} + \begin{bmatrix} e_{111} \\ e_{112} \\ e_{121} \\ e_{122} \\ e_{211} \\ e_{212} \\ e_{221} \\ e_{222} \end{bmatrix}$$

Then

$$\mathbf{D} = \left[\begin{array}{cc|cccc|cccc}
 \sigma_g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \sigma_g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gt_1}^2 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gt_2}^2 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gt_1}^2 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gt_2}^2
 \end{array} \right]$$

and

$$\mathbf{R} = \sigma_e^2 \mathbf{I}_8$$

where \mathbf{I}_8 denotes an (8×8) identity matrix. Therefore, \mathbf{V} can be written in block form as

$$\mathbf{V} = \left[\begin{array}{cc|cc}
 \mathbf{A} & \mathbf{B} & \mathbf{0} & \mathbf{0} \\
 \mathbf{B} & \mathbf{A} & \mathbf{0} & \mathbf{0} \\
 \hline
 \mathbf{0} & \mathbf{0} & \mathbf{A} & \mathbf{B} \\
 \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{A}
 \end{array} \right]$$

where

$$\mathbf{A} = \begin{bmatrix} \sigma_g^2 + \sigma_{gt_1}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gt_2}^2 + \sigma_W^2 \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \sigma_g^2 + \sigma_{gt_1}^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 + \sigma_{gt_2}^2 \end{bmatrix}$$

with $\sigma_W^2 = \sigma_s^2 + \sigma_e^2$. We see based on \mathbf{V} that observations from different groups are uncorrelated (note the (4×4) block of zeros in the upper right and lower left of \mathbf{V}). Note that \mathbf{A} represents the covariance matrix for

observations coming from a single individual (across the two time points), while \mathbf{B} represents the covariance matrix for observations from different individuals in the same group (across the two time points). Hence, the correlation between observations from individuals within the same group is given by

$$ICC_1 = \frac{\sigma_g^2 + \sigma_{gt_1}^2}{\sigma_g^2 + \sigma_{gt_1}^2 + \sigma_W^2} \quad (\text{at time point 1})$$

$$ICC_2 = \frac{\sigma_g^2 + \sigma_{gt_2}^2}{\sigma_g^2 + \sigma_{gt_2}^2 + \sigma_W^2} \quad (\text{at time point 2})$$

THE MOTIVATING EXAMPLE

Now consider the motivating example discussed by Cohen and Doveh (this volume). Here $Y_{i(j)c}$ denotes an observation from subject $j = 1, 2, 3$ in group $i = 1, \dots, 40$ and condition $c = 1, \dots, 4$. The model is exactly of the same form as the previous one, except that the repeated factor “condition” now has four levels (i.e., low interdependence task at pre-test, low interdependence task at post-test, high interdependence task at pre-test, high interdependence task at post-test). Adapting the notation slightly to reflect this, the model can then be specified as follows:

$$Y_{i(j)c} = \mu + (con)_c + (grp)_i + (subj)_{j(i)} + (grp \times con)_{ic} + e_{i(j)c}$$

The variances of the group and subject effects will still be denoted by σ_g^2 and σ_s^2 , respectively, while the variances of the group by condition interaction effect will be denoted, in the most general case, by $\sigma_{gc_1}^2, \dots, \sigma_{gc_4}^2$ (for the four measurement occasions).

We will again write out the model in GLMM notation. However, we will do so only for the three subjects in the first group. As noted above, observations from different groups are independent and therefore looking at a single group is sufficient for our purposes. Thus, for the first group, we write

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{113} \\ Y_{114} \\ \hline Y_{121} \\ Y_{122} \\ Y_{123} \\ Y_{124} \\ \hline Y_{131} \\ Y_{132} \\ Y_{133} \\ Y_{134} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ s_{11} \\ s_{21} \\ s_{31} \\ gc_{11} \\ gc_{12} \\ gc_{13} \\ gc_{14} \end{bmatrix} + \begin{bmatrix} e_{111} \\ e_{112} \\ e_{113} \\ e_{114} \\ \hline e_{121} \\ e_{122} \\ e_{123} \\ e_{124} \\ \hline e_{131} \\ e_{132} \\ e_{133} \\ e_{134} \end{bmatrix}$$

Then

$$D = \begin{bmatrix} \sigma_g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{gc_1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{gc_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gc_3}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gc_4}^2 \end{bmatrix}$$

and

$$R = \sigma_e^2 I_{12}$$

where I_{12} denotes a (12×12) identity matrix. The form of V for the first (or any other) group is now given by

$$V = \begin{bmatrix} A & B & B \\ B & A & B \\ B & B & A \end{bmatrix}$$

where

$$A = \begin{bmatrix} \sigma_g^2 + \sigma_{gc_1}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_2}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_3}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_4}^2 + \sigma_W^2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} \sigma_g^2 + \sigma_{gc_1}^2 & \sigma_g^2 & \sigma_g^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 + \sigma_{gc_2}^2 & \sigma_g^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 & \sigma_g^2 + \sigma_{gc_3}^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 & \sigma_g^2 & \sigma_g^2 + \sigma_{gc_4}^2 \end{bmatrix}$$

with $\sigma_W^2 = \sigma_s^2 + \sigma_e^2$. As before, A is the covariance matrix for observations coming from a single individual (across the four conditions), while B is the covariance matrix for observations from two different individuals in the

same group (across the four conditions). Therefore, observations from individuals within the same group are correlated as follows:

$$ICC_1 = \frac{\sigma_g^2 + \sigma_{gc_1}^2}{\sigma_g^2 + \sigma_{gc_1}^2 + \sigma_W^2} \quad (\text{low interdependence task at pre-test})$$

$$ICC_2 = \frac{\sigma_g^2 + \sigma_{gc_2}^2}{\sigma_g^2 + \sigma_{gc_2}^2 + \sigma_W^2} \quad (\text{low interdependence task at post-test})$$

$$ICC_3 = \frac{\sigma_g^2 + \sigma_{gc_3}^2}{\sigma_g^2 + \sigma_{gc_3}^2 + \sigma_W^2} \quad (\text{high interdependence task at pre-test})$$

$$ICC_4 = \frac{\sigma_g^2 + \sigma_{gc_4}^2}{\sigma_g^2 + \sigma_{gc_4}^2 + \sigma_W^2} \quad (\text{high interdependence task at post-test})$$

Changes in group homogeneity can now be assessed by examining whether the ICC changes over time and/or experimental tasks.

Specification of all Possible Subset Models

The model defined in the previous section was not actually considered by Cohen and Doveh (this volume) in their analysis. Instead, using a bottom-up approach, Cohen and Doveh examined a series of simpler models, in which at least some of the four variance components for the interaction effect (i.e., $\sigma_{gc_1}^2, \dots, \sigma_{gc_4}^2$) were constrained to be equal to each other – and by implication, so were the corresponding ICCs (e.g., $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ implies that $ICC_1 = ICC_2$). We can in fact consider a large number of possible models, depending on the number of constraints we want to impose on the variance components. Table 1 provides a listing of all possible models.

For example, in model 8, $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ and $\sigma_{gc_3}^2 = \sigma_{gc_4}^2$ and consequently $ICC_1 = ICC_2$ and $ICC_3 = ICC_4$. In other words, observations from individuals in the same group show the same degree of correlation at the pre- and post-test, but the correlation may differ depending on task type. Hence, this model implies no change in group homogeneity across the two time points, but possibly across the two task types.

The various constraints are imposed on model 1 by simple changes to the **D** matrix. For example, for model 8, we let

Table 1. Listing of all Possible Models Depending on the Equality Constraints among the Interaction Variances.

	$\sigma_{gc_1}^2$	$\sigma_{gc_2}^2$	$\sigma_{gc_3}^2$	$\sigma_{gc_4}^2$	No. of Free Parameters
Model 1	a	b	c	d	4
Model 2	a	a	b	c	3
Model 3	a	b	a	c	3
Model 4	a	b	c	a	3
Model 5	a	b	b	c	3
Model 6	a	b	c	b	3
Model 7	a	b	c	c	3
Model 8	a	a	b	b	2
Model 9	a	b	a	b	2
Model 10	a	b	b	a	2
Model 11	a	a	a	b	2
Model 12	a	a	b	a	2
Model 13	a	b	a	a	2
Model 14	b	a	a	a	2
Model 15	a	a	a	a	1

Note: Variance components with the same letter are constrained to be equal to each other.

$$D = \begin{bmatrix} \sigma_g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_s^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{gc_a}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{gc_a}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gc_b}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{gc_b}^2 \end{bmatrix}$$

and therefore

$$A = \begin{bmatrix} \sigma_g^2 + \sigma_{gc_a}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_a}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_b}^2 + \sigma_W^2 & \sigma_g^2 + \sigma_s^2 \\ \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_s^2 & \sigma_g^2 + \sigma_{gc_b}^2 + \sigma_W^2 \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \sigma_g^2 + \sigma_{gc_a}^2 & \sigma_g^2 & \sigma_g^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 + \sigma_{gc_a}^2 & \sigma_g^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 & \sigma_g^2 + \sigma_{gc_b}^2 & \sigma_g^2 \\ \sigma_g^2 & \sigma_g^2 & \sigma_g^2 & \sigma_g^2 + \sigma_{gc_b}^2 \end{bmatrix}$$

from which follows that

$$ICC_1 = ICC_2 = \frac{\sigma_g^2 + \sigma_{gc_a}^2}{\sigma_g^2 + \sigma_{gc_a}^2 + \sigma_W^2}$$

(low interdependence task at pre- and post-test)

$$ICC_3 = ICC_4 = \frac{\sigma_g^2 + \sigma_{gc_b}^2}{\sigma_g^2 + \sigma_{gc_b}^2 + \sigma_W^2}$$

(high interdependence task at pre- and post-test)

Therefore, the various models specify all the different ways how group homogeneity may change over time and/or experimental tasks.

These 15 possible models can be classified into five groups, as a function of the number of free parameters being estimated: (a) the full model (model 1 in the table) including four ICCs (one for each combination of task by time), (b) six subset models (models 2–7 in the table) with only three ICCs (assuming that two of the four ICCs are equal), (c) three subset models (models 8–10) with only two ICCs (assuming that two pairs of ICCs are equal), (d) four subset models (models 11–14) with only two ICCs (assuming that three of the four ICCs are equal), and (e) a one-parameter model (model 15) assuming equality of all four ICCs. Also note that this hierarchy involves models that are properly nested within each other (e.g., all the models are properly nested within model 1 and the constraints for model 8 are a proper subset of those for model 2) and models that are not properly nested within each other (e.g., models 9 and 10 are not nested in model 2).

It is possible to fit each of the 15 models using the same approach and software (SAS code for fitting the 15 models is given in the appendix), although clearly some models make more intuitive sense than others. For example, the constraints defining model 8 (equal ICCs under low interdependence at both times and equal ICCs under high interdependence at both times) and model 9 (equal ICCs at the pre-test under both levels of interdependence and equal ICCs at the post-test under both levels of

Table 2. Estimates of Variance Components and Standard Errors for Model 1.

Variance Component		Group	Estimate	Standard Error
Id(grp)	σ_s^2		86.3716	22.9022
grp	σ_g^2		-1.2464	17.6657
grp * cond	σ_{gc1}^2	1	59.0292	36.0487
grp * cond	σ_{gc2}^2	2	82.6210	40.4430
grp * cond	σ_{gc3}^2	3	135.28	52.6998
grp * cond	σ_{gc4}^2	4	69.2271	38.1404
Residual	σ_e^2		216.91	19.9877

interdependence) are naturally appealing, while the constraints of model 10 make less sense and are less compelling.

One particular aspect of Cohen and Doveh's approach (this volume) calls for special attention. While Cohen and Doveh are examining a variety of simpler models, they are actually considering only subsets of the entire dataset that are relevant for establishing the homogeneity of various variance components. For example, in the first step, Cohen and Doveh examine whether $\sigma_{gc1}^2 = \sigma_{gc2}^2$. Consequently, only the pre- and post-test data from the low interdependence task are used. This is similar to comparing model 1 with model 2, but instead of including all the data and leaving σ_{gc3}^2 and σ_{gc4}^2 free to vary (as model 2 would allow), the data for the high interdependence task are excluded. Similarly, only subsets of the entire dataset are used when examining whether $\sigma_{gc1}^2 = \sigma_{gc3}^2$ and whether $\sigma_{gc1}^2 = \sigma_{gc2}^2 = \sigma_{gc3}^2$.

An advantage of using the full dataset is that the remaining variance components (i.e., σ_g^2 , σ_s^2 and σ_e^2) are estimated more efficiently, thereby also leading to more efficient ICC estimates. For example, Table 12 in Cohen and Doveh (this volume) shows the estimated variance components and associated standard errors when only using the pre- and post-test data from the low interdependence task (and allowing the ICCs for the pre- and post-test data to differ). When using the full dataset and fitting model 1 (which estimates all four ICCs without imposing any constraints), then the estimates as shown in Table 2 are obtained. Comparing these estimates with those given by Cohen and Doveh, we note that the standard errors for $\hat{\sigma}_s^2$, $\hat{\sigma}_g^2$ and $\hat{\sigma}_e^2$ are smaller when using the full dataset.

However, considering that the residual variances were later found to be heterogeneous (i.e., σ_e^2 also appears to depend on the condition level), it is

Table 3. ICCs under all Possible Models Depending on the Equality Constraints among the Interaction Variances.

		ICC_1	ICC_2	ICC_3	ICC_4	$-2ll_R$	AIC
Model 1	abcd	0.16	0.21	0.31	0.18	4105.3	4119.3
Model 2	aabc	0.19	0.19	0.31	0.18	4105.6	4117.6
Model 3	abac	0.24	0.23	0.24	0.17	4106.9	4118.9
Model 4	abca	0.17	0.21	0.31	0.17	4105.4	4117.4
Model 5	abbc	0.17	0.26	0.26	0.18	4106.0	4118.0
Model 6	abcb	0.16	0.20	0.31	0.20	4105.4	4117.4
Model 7	abcc	0.16	0.21	0.25	0.25	4106.6	4118.6
Model 8	aabb	0.19	0.19	0.25	0.25	4106.8	4116.8
Model 9	abab	0.24	0.20	0.24	0.20	4107.0	4117.0
Model 10	abba	0.17	0.26	0.26	0.17	4106.0	4116.0
Model 11	aaab	0.24	0.24	0.24	0.17	4106.9	4116.9
Model 12	aaba	0.19	0.19	0.31	0.19	4105.6	4115.6
Model 13	abaa	0.22	0.22	0.22	0.22	4107.4	4117.4
Model 14	baaa	0.16	0.24	0.24	0.24	4106.7	4116.7
Model 15	aaaa	0.22	0.22	0.22	0.22	4107.4	4115.4

Note: $-2ll_R = -2$ times the restricted log-likelihood; $AIC =$ Akaike's information criterion.

understandable why Cohen and Doveh (this volume) chose to focus on subsets of the data initially. By focusing on subsets in which the residual variances appear to be roughly homogeneous, the problem of having to model heterogeneous residual variances is circumvented at first.

Nevertheless, we have some reservations about the approach chosen by Cohen and Doveh (this volume). They first test whether $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ (the first pair-wise comparison) and $\sigma_{gc_1}^2 = \sigma_{gc_3}^2$ (the second pair-wise comparison) and subsequently test whether $\sigma_{gc_1}^2 = \sigma_{gc_2}^2 = \sigma_{gc_3}^2$ (the three-way comparison). With this approach, one may reject $\sigma_{gc_1}^2 = \sigma_{gc_2}^2 = \sigma_{gc_3}^2$ after having previously established that $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ and $\sigma_{gc_1}^2 = \sigma_{gc_3}^2$. Such inconsistent and intransitive patterns have long been a thorny issue for multiple comparisons procedures (e.g., Shaffer, 1995). By using Table 1 to choose a sequence of properly nested models, one can avoid such contradictions.

Model Selection

In Table 3, we are providing a listing of the ICCs for the four conditions under the 15 models specified in Table 1. Also given are the value of -2 times the restricted log-likelihood ($-2ll_R$) and AIC for each model. The AIC

is here defined as

$$AIC = -2ll_R + 2q$$

where q is number of variance components in the model (Wolfinger, 1993). Smaller values of $-2ll_R$ indicate a better fitting model, but the fit of a model can always be improved, at least marginally, by increasing the complexity of the model. The AIC is a measure of model fit that includes a penalty for increasing the model complexity. Therefore, models with a small AIC strike a good balance between fit and complexity (Bozdogan, 1987).

Using the deviance test, we find that model 15 (which constrains all ICCs equal to each other) fits just as well as model 1 (which leaves all four ICCs unconstrained): $\Delta = 4107.4 - 4105.3 = 2.1$, which under the null hypothesis follows a χ^2 distribution with 3 degrees of freedom and therefore is far from being significant ($p = 0.55$). In fact, considering that the smallest possible critical value for model comparisons is 3.84 (for a χ^2 distribution with $df = 1$, i.e., for imposing a single constraint), any model comparison is going to be not significant (for the models in Table 3, the largest value of Δ is 2.1). Also note that the AIC of model 15 is the smallest, indicating that this model strikes the best balance between fit and complexity. Therefore, while the estimates from model 1 seem to indicate approximately homogeneous ICCs in conditions 1, 2, and 4 and a slightly higher ICC in condition 3, the model comparison shows that model 15 fits just as well, indicating homogeneous ICCs across the four conditions.

Models with Heterogeneous Residual Variances

However, the real “gist” of the example is missed by the analysis so far. As was mentioned before, the residual variance (σ_e^2) may also depend on the condition level. Accordingly, we can consider an even more general model, where \mathbf{R} is now given by three diagonal blocks of

$$\begin{bmatrix} \sigma_{e_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{e_2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{e_3}^2 & 0 \\ 0 & 0 & 0 & \sigma_{e_4}^2 \end{bmatrix}$$

instead of $\mathbf{R} = \sigma_e^2 \mathbf{I}_{12}$. Now \mathbf{V} has the same structure as before, except that the $\sigma_{W_i}^2$ values along the diagonal of \mathbf{A} are given by $\sigma_{W_1}^2 = \sigma_s^2 + \sigma_{e_1}^2, \dots, \sigma_{W_4}^2 = \sigma_s^2 + \sigma_{e_4}^2$, depending on the condition level. The four ICCs

are then equal to:

$$ICC_1 = \frac{\sigma_g^2 + \sigma_{gc_1}^2}{\sigma_g^2 + \sigma_{gc_1}^2 + \sigma_{W_1}^2} \quad (\text{low interdependence task at pre-test})$$

$$ICC_2 = \frac{\sigma_g^2 + \sigma_{gc_2}^2}{\sigma_g^2 + \sigma_{gc_2}^2 + \sigma_{W_2}^2} \quad (\text{low interdependence task at post-test})$$

$$ICC_3 = \frac{\sigma_g^2 + \sigma_{gc_3}^2}{\sigma_g^2 + \sigma_{gc_3}^2 + \sigma_{W_3}^2} \quad (\text{high interdependence task at pre-test})$$

$$ICC_4 = \frac{\sigma_g^2 + \sigma_{gc_4}^2}{\sigma_g^2 + \sigma_{gc_4}^2 + \sigma_{W_4}^2} \quad (\text{high interdependence task at post-test})$$

Analogously to the interaction variances, we can let all or only subsets of the residual variances be heterogeneous. We will make the simplifying “matching homogeneity” assumption that if $\sigma_{gc_i}^2 = \sigma_{gc_j}^2$, then $\sigma_{e_i}^2 = \sigma_{e_j}^2$ for all pairs i, j . We do this for two reasons. First of all, without this assumption, the number of possible models would amount to $15 \times 15 = 225$ (since we could combine 15 different types of constraints on the residual variances with each of the 15 models in Table 1). The simplifying assumption reduces the total number of models down to 15. However, more importantly, if we can simultaneously show that $\sigma_{gc_i}^2 = \sigma_{gc_j}^2$ and $\sigma_{e_i}^2 = \sigma_{e_j}^2$, then this again implies $ICC_i = ICC_j$. In other words, the assumption allows us to establish the homogeneity of various ICCs through corresponding reductions in the model complexity (more about this issue will be said below).

Model Selection

Results for fitting each of these models with heterogeneous residual variances are given in Table 4 (SAS code can again be found in the appendix) and a comparison with Table 3 demonstrates that the results generally differ quite substantially from those obtained under the assumption of homogeneous residual variances (except for model 15(h) in Table 4, which is identical to model 15 in Table 3 by definition).

Model 1(h) (the most general model with heterogeneous residual variances) fits significantly better than model 1 (the most general model with homogeneous residual variances): $\Delta = 4105.3 - 4090.1 = 15.2$, which is

Table 4. ICCs under all Possible Models Depending on the Equality Constraints among the Interaction and Residual Variances.

		ICC_1	ICC_2	ICC_3	ICC_4	$-2ll_R$	AIC
Model 1(h)	abcd	0.05	0.23	0.25	0.35	4090.1	4110.1
Model 2(h)	aabc	0.14	0.14	0.26	0.34	4093.3	4109.3
Model 3(h)	abac	0.16	0.24	0.16	0.33	4092.2	4108.2
Model 4(h)	abca	0.19	0.23	0.25	0.19	4103.6	4119.6
Model 5(h)	abbc	0.06	0.24	0.24	0.34	4091.6	4107.6
Model 6(h)	abcb	0.07	0.28	0.24	0.28	4095.3	4111.3
Model 7(h)	abcc	0.07	0.22	0.30	0.30	4101.9	4117.9
Model 8(h)	aabb	0.14	0.14	0.30	0.30	4104.1	4116.1
Model 9(h)	abab	0.16	0.28	0.16	0.28	4097.2	4109.2
Model 10(h)	abba	0.19	0.24	0.24	0.19	4105.3	4117.3
Model 11(h)	aaab	0.19	0.19	0.19	0.33	4094.6	4106.6
Model 12(h)	aaba	0.20	0.20	0.25	0.20	4103.8	4115.8
Model 13(h)	abaa	0.22	0.24	0.22	0.22	4107.2	4119.2
Model 14(h)	baaa	0.07	0.27	0.27	0.27	4102.4	4114.4
Model 15(h)	aaaa	0.22	0.22	0.22	0.22	4107.4	4115.4

See Note in Table 3.

distributed χ^2 with 3 degrees of freedom under the null hypothesis that $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \sigma_{e_3}^2 = \sigma_{e_4}^2$ and therefore $p = 0.002$. Also, the AIC of model 1(h) clearly falls below that of model 1, indicating that the increase in model complexity due to heterogeneous residual variances is outweighed by an even larger increase in model fit. Finally, note that the resulting ICCs under model 1(h) provide a very different picture than those obtained under model 1: an ICC of almost zero under condition 1, similar ICCs under conditions 2 and 3, and a slightly higher ICC in condition 4.

Having determined that heterogeneous residual variances are indicated, we can again use a hierarchical model selection approach (this time using Table 4) to find a more parsimonious model. Analogously to the approach taken by Cohen and Doveh (this volume), we first compare model 1(h) against model 2(h) and find that the simpler model fits just as well ($\Delta = 4093.3 - 4090.1 = 3.2$, $df = 2$, $p = 0.36$). At this point, we can consider models 8(h), 11(h), and 12(h), which are all properly nested within model 2(h). It turns out that model 11(h) fits just as well as model 2(h) ($\Delta = 4094.6 - 4093.3 = 1.3$, $df = 2$, $p = 0.52$), while models 8(h) and 12(h) fit significantly worse ($p = 0.005$ in both comparisons). Finally, model 15(h) fits significantly worse than model 11(h) ($p = 0.002$) and therefore the hierarchical selection approach stops at this point. Note also that model 1

1(h) has the smallest AIC value of all the models considered. In essence, model 1 1(h) implies that group homogeneity does not change from pre- to post-test in the low interdependence task, but increases over time in the high-interdependence task.

Additional Complexity due to Heterogeneous Residual Variances

It is important to point out that this hierarchical selection approach focuses at each step on the question of model fit, but does not directly test the equality of the various ICCs. Specifically, as Cohen and Doveh (this volume) note, the introduction of heterogeneous residual variances complicates matters with respect to directly testing the equality of ICCs. For homogeneous residual variances, we could test whether two ICCs are equal to each other by simply testing the equality of the corresponding interaction variances. This no longer works when the residual variances are heterogeneous. For example, $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ no longer implies that $ICC_1 = ICC_2$. At the same time, simultaneously testing whether $\sigma_{gc_1}^2 = \sigma_{gc_2}^2$ and $\sigma_{W_1}^2 = \sigma_{W_2}^2$, (to determine whether $ICC_1 = ICC_2$) is too restrictive, because it is possible that $\sigma_{gc_1}^2 \neq \sigma_{gc_2}^2$ and $\sigma_{W_1}^2 \neq \sigma_{W_2}^2$, but $ICC_1 = ICC_2$. For example:

$$ICC_1 = \frac{\sigma_g^2 + \sigma_{gc_1}^2}{\sigma_g^2 + \sigma_{gc_1}^2 + \sigma_{W_1}^2} = \frac{100 + 40}{100 + 40 + 420} = 0.25$$

$$ICC_2 = \frac{\sigma_g^2 + \sigma_{gc_2}^2}{\sigma_g^2 + \sigma_{gc_2}^2 + \sigma_{W_2}^2} = \frac{100 + 80}{100 + 80 + 540} = 0.25$$

Similarly, it might be the case in the motivating example that $\sigma_{gc_1}^2 = \sigma_{gc_2}^2 = \sigma_{gc_3}^2 \neq \sigma_{gc_4}^2$ and $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \sigma_{e_3}^2 \neq \sigma_{e_4}^2$, but $ICC_1 = \dots = ICC_4$. In other words, all four ICCs may be homogeneous, despite the fact that model 11(h) fits significantly better than model 15(h).

As Cohen and Doveh (this volume) point out, there is currently no straightforward way of directly testing the equality of ICCs in this case. However, as long as we can reduce the model complexity by simultaneously imposing the same constraints on the residual and the interaction variances, we suggest that this additional complexity is of minor concern. Specifically, if the model complexity can be reduced through such simultaneous constraints (e.g., when going from model 1(h) to model 2(h) and from model 2(h) to model 11(h)), then we are not only finding models whose fit is not significantly worse than that of the more complex models, but we are also demonstrating the homogeneity of subsets of ICCs.

However, once the model complexity cannot be reduced further by simultaneously imposing constraints on the residual and the interaction variances, then we must proceed in a different manner. In particular, we can conduct two separate tests, as Cohen and Doveh (this volume) do. It turns out that model 11(h) can be further simplified, since constraining all four interaction variances equal to each other (i.e., $\sigma_{gc_1}^2 = \sigma_{gc_2}^2 = \sigma_{gc_3}^2 = \sigma_{gc_4}^2$) does not reduce the fit of the model substantially ($p = 0.654$), while a test of $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \sigma_{e_3}^2 = \sigma_{e_4}^2$ does turn out to be significant ($p < 0.001$). Consequently, the final ICCs are given by

$$\begin{aligned} ICC_1 = ICC_2 = ICC_3 \\ &= \frac{\sigma_g^2 + \sigma_{gc_a}^2}{\sigma_g^2 + \sigma_{gc_a}^2 + \sigma_{W_a}^2} \\ &\quad \text{(low interdependence task at pre- and post-test} \\ &\quad \text{and high interdependence task at pre-test)} \end{aligned}$$

$$ICC_4 = \frac{\sigma_g^2 + \sigma_{gc_a}^2}{\sigma_g^2 + \sigma_{gc_a}^2 + \sigma_{W_b}^2} \quad \text{(high interdependence task at post-test)}$$

We therefore have indirectly determined that the triplet of homogeneous ICCs is significantly different from ICC_4 .

CONCLUSION

Cohen and Doveh (this volume) should be congratulated for their excellent discussion of the topic of comparing dependent ICCs. In the present chapter we have shown explicitly how the motivating example can be extended and framed in the context of the general linear mixed-effects model. The suggested approach stresses the representation of all the hypotheses being tested as explicit model comparisons and can be construed as a search process whose goal is, essentially, the identification of the best-fitting model. This approach leads naturally to a consideration of all possible models (e.g., Table 1) and highlights the commonalities as well as the key differences between them. By using the same standard parameter estimation technique on all the data in the sample for all the models being considered, this approach is (a) more powerful and efficient, (b) minimizes the chances of inconsistencies in inferences across tests, and (c) provides the researchers with a deeper and more complete understanding of the results.

In this particular case, we selected the same final model and arrived at the same conclusions with respect to the motivating example as Cohen and Doveh (this volume) did in their analysis, which is nice and reassuring. However, this is not guaranteed to be the case in all circumstances. We therefore consider it important to take into account the approach discussed in our commentary when examining data of the type described herein.

REFERENCES

- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.
- Dayton, C. M. (2003). Information criteria for pairwise comparisons. *Psychological Methods*, *8*, 61–71.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NY: Lawrence Erlbaum.
- Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Chicago: Irwin.
- Searle, S. R., Casella, G., & McCulloch, C. E. (1992). *Variance components*. New York: Wiley.
- Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, *46*, 561–584.
- Wolfinger, R. (1993). Covariance structure selection in general mixed models. *Communications in Statistics: Simulation and Computation*, *22*, 1079–1106.

APPENDIX

SAS code for fitting the various models presented in this commentary are given below. Assume that the data for the motivating example are contained in a plain text file (“data.txt”) with the following structure:

1	1	low1	44
1	1	low2	30
1	1	high1	17
1	1	high2	45
...			

The first number is the id number of the subject (1, ..., 120), the second value indicates the group number (1, ..., 40), the third variable indicates the condition level (low1, low2, high1, and high2), and the final value indicates the subject's response. Then the 15 models with homogeneous residual variances can be fit with the following code:

```
data example;
  infile data.txt;
  input id grp cond $ y;
run;

data example;
  set example;
  if cond = 'low1' then cat = 'a';
  if cond = 'low2' then cat = 'b';
  if cond = 'high1' then cat = 'c';
  if cond = 'high2' then cat = 'd';
run;

proc mixed data = example covtest nobound;
  class grp cond id cat;
  model y = cond;
  random int /subject = id(grp);
  random int /subject = grp vcorr;
  random int /subject = cond*grp group = cat;
run;
```

By adjusting the letters in the various if-statements according to Table 1, one can impose the various equality constraints (the code as given above would fit model 1). In order to fit the 15 models with heterogeneous residual variances, one simply needs to add

```
repeated /type = vc subject = grp group = cat;
```

to the proc mixed step (at any point below the line starting with “model”).