Behavior Research Methods (2023) 56:2013–2032
https://doi.org/10.3758/s13428-023-02133-1

Alternative covariance structures in mixed-effects models: Addressing intra- and inter-individual heterogeneity

Shelley A. Blozis[1] · **Madeline Craft**[1]

Accepted: 21 April 2023 / Published online: 25 May 2023

Abstract

Mixed-effects models for repeated measures and longitudinal data include random coefficients that are unique to the individual, and thus permit subject-specific growth trajectories, as well as direct study of how the coefficients of a growth function vary as a function of covariates. Although applications of these models often assume homogeneity of the within-subject residual variance that characterizes within-person variation after accounting for systematic change and the variances of the random coefficients of a growth model that quantify individual differences in aspects of change, alternative covariance structures can be considered. These include allowing for serial correlations between the within-subject residuals to account for dependencies in data that remain after fitting a particular growth model or specifying the within-subject residual variance to be a function of covariates or a random subject effect to address between-subject heterogeneity due to unmeasured influences. Further, the variances of the random coefficients can be functions of covariates to relax the assumption that these variances are constant across subjects and to allow for the study of determinants of these sources of variation. In this paper, we consider combinations of these structures that permit flexibility in how mixed-effects models are specified to understand within- and between-subject variation in repeated measures and longitudinal data. Data from three learning studies are analyzed using these different specifications of mixed-effects models.

Keywords Autocorrelation · Procedural learning · Response latency · Working memory · Mixed-effects location scale models · Nonlinear mixed-effects models

Repeated measures and longitudinal data are essential in studies of growth or change in human behavior. Data collected across repeated occasions, whether the occasions are closely spaced (e.g., studies of human learning or ecological momentary assessments to understand variation in mood) or span across a relatively long period of time (e.g., longitudinal studies of human growth and development), are necessary to understand if, when, or how behaviors change and how other variables might influence the process.

Mixed-effects models are widely applied for the analysis of repeated measures and longitudinal data. These models emphasize the individual by specifying a growth model at the subject level and provide a framework to study within- and between-subject variation in measured behaviors. In specifying a model, a function, typically assumed to be common to all individuals, is selected to describe the response, but one or more of the function coefficients are subject-specific to allow for individual differences in certain aspects of change or development. In this way, the subject-specific model can be used to account for variation in scores within individuals, and the subject-specific coefficients permit the study of how individuals differ in aspects of change. An appealing quality of these models is that the coefficients of a function can vary according to covariates at the subject level to study how specific aspects of change or development depend on covariates.

Careful selection of a growth function that effectively summarizes a large collection of data by a relatively small number of parameters can aid in the analysis and interpretation of repeated measures and longitudinal data, especially when the form of change is complex (Cudeck & Harring, 2007). Further, if the chosen function is effective in capturing variation within individuals, then the occasion-specific residuals will be independent within subjects. In other words, it is typically reasonable to assume that scores within subjects are correlated, and a mixed-effects model accounts

✉ Shelley A. Blozis
sablozis@ucdavis.edu

[1] Department of Psychology, University of California, Davis, Davis, California, USA

for these correlations with the random coefficients of the growth model. If, however, a function is ineffective in fully accounting for this variation, some dependencies between residuals will remain.

Previous work has documented the impact on model fit and statistical inference of poorly specified residual covariance structures at the first level of a mixed-effects model. It is well known that applications of mixed-effects models that erroneously assume that the residuals are independent with constant variance across occasions have consequences for the estimates of the variances of the random coefficients of the subject-level model (Baek et al., 2020; Baek & Ferron, 2020; Blozis & Harring, 2021; Chi & Reinsel, 1989; Ferron et al., 2002; Harring & Blozis, 2014; Joo et al., 2019; Sivo et al., 2005). Thus, in fitting a mixed-effects model, the analyst faces a need to balance parsimony and interpretability in selecting a growth model with the need to adequately address dependencies of scores within subjects.

A special class of mixed-effects models, known as mixed-effects location scale models, have expanded the ways in which to model intra- and inter-individual differences by including submodels for the variance of the residuals at the first level of the model, as well as for the variances of the random coefficients at the second level (for three-level models, see Lin et al., 2018; Nestler et al., 2018). Importantly, these extensions provide a framework to relax assumptions of homogeneity of the covariance structure at each model level, in addition to providing a means to study the determinants of these sources of variation in data. These models grew from the need to model repeated-measures data that generally did not involve growth or change, and the residuals at level 1 were assumed to be independent (see Hedeker et al., 2008), but their application to growth and change processes (e.g., McNeish, 2021; Williams et al., 2019) have not consistently incorporated this earlier literature that has emphasized the importance of considering alternative residual covariance structures when the assumption of independence between the level-1 residuals is not tenable. Shown to be useful in applications of linear growth models (Nestler, 2021, 2022), this paper joins these two methodological areas within the context of fitting nonlinear mixed-effects models, thus broadening their use in problems involving relatively complex forms of change.

To accomplish this, data from three learning experiments are utilized to motivate researchers to expand their thinking about how they might formulate the covariance structure of a mixed-effects model to study within- and between-individual differences in repeated measures that follow complex forms of change. We first develop a general framework to increase options in how one specifies the covariance structure of a mixed-effects model with special attention to how covariates may be incorporated into the structure. The examples are then presented. We develop syntax for using maximum likelihood (ML) estimation of models using SAS PROC NLMIXED by expanding on the developments of Harring and Blozis (2014). A discussion follows with implications and directions for future research.

Mixed-effects models for repeated measures and longitudinal data

Resources on mixed-effects models are numerous. Among the many published materials on mixed-effects models are books that serve a range of purposes, with some providing a general resource (Wu, 2010) or focus on a particular software program (Rabe-Hesketh & Skrondal, 2012) or area of application (Brown & Prescott, 2015), and others emphasizing specific topics, such as nonlinear mixed-effects models (Fiedler-Kelly & Owen, 2014) or autoregressive mixed-effects models (Funatogawa & Funatogawa, 2018). We assume readers are familiar with common formulations of a mixed-effects model, including linear and nonlinear models. We give a brief description of the model that serves as a starting point for our developments here.

Let y_{ij} be the observed measure for individual i at time j, where $i = 1,..., N$ and $j = 1,..., n_i$, letting N denote the number of subjects and n_i the number of measures for individual i. Let t_{ij} be the time when y_{ij} was observed. A mixed-effects model for y_{ij} is

$$y_{ij} = f\left(t_{ij}, \mathbf{X}_{ij}, \boldsymbol{\theta}_i, \mathbf{W}_i, \boldsymbol{\gamma}\right) + \varepsilon_{ij} \quad (1)$$

where $f(\cdot)$ is a function (e.g., linear, quadratic, logistic) assumed to characterize the growth or developmental trend of the response. In (1), y_{ij} is a function of t_{ij}, a set of covariates $\mathbf{X}_{ij}$ (that usually includes 1 for the intercept of the model) that vary with y_{ij}, a set of subject-specific coefficients $\boldsymbol{\theta}_i$ that link $\mathbf{X}_{ij}$ and y_{ij}, a set of between-subject covariates $\mathbf{W}_i$, a set of fixed coefficients $\boldsymbol{\gamma}$ that link $\mathbf{W}_i$ and y_{ij}, and a time- and subject-specific residual ε_{ij}. Coefficients in $\boldsymbol{\theta}_i$ can be fixed, random, or the sum of a fixed and a random effect. That is, for the qth coefficient in $\boldsymbol{\theta}_i$, $\theta_{qi} = \alpha_q + u_{qi}$, where α_q is the fixed effect and u_{qi} is the corresponding random effect. The random effects are assumed to be independently and identically distributed (i.i.d.) as normal across subjects with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}$, where the dimensions of $\boldsymbol{\Phi}$ depend on the number of random effects in $\boldsymbol{\theta}_i$. The residual ε_{ij} is the response not accounted for by the subject-specific model given by $f\left(t_{ij}, \mathbf{X}_{ij}, \boldsymbol{\theta}_i, \mathbf{W}_i, \boldsymbol{\gamma}\right)$. The set of residuals $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},...,\varepsilon_{in_i})'$ is assumed to be i.i.d. normal across subjects with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Theta}_\varepsilon$, where the dimensions of $\boldsymbol{\Theta}_\varepsilon$ depends on n_i.

Within-subject covariance structure

The residual covariance matrix characterizes the variation and between-occasion covariation of the deviations of observations from their expected values. For a growth model that fully accounts for the within-subject dependencies of scores between occasions and given that the residual variances are equal across time, a simple structure for the covariance matrix is appropriate: $\mathbf{\Theta}_{\varepsilon} = \sigma^2_{\varepsilon}\mathbf{I}_{n_i}$. Other options ought to be considered, however, if within-subject dependencies between scores remain, even after accounting for covariates, or if the variances of the residuals are not equal across occasions. The assumption of homogeneity of the residual covariance structure between subjects can be relaxed by specifying a model in which the covariance structure differs between groups, such as by allowing the variance of a simple covariance structure to differ between groups: $\mathbf{\Theta}_{\varepsilon k} = \sigma^2_{\varepsilon k}\mathbf{I}_{n_{ik}}$, where k denotes group membership. The variance itself could also be a function of measured covariates (Stroup et al., 2018), similar to linear regression models that use an exponential function to model the residual variance to address heterogeneity (Aitkin, 1987; Cook & Weisberg, 1983; Harvey, 1976; Carroll & Ruppert, 1988). In such cases, between-subject heterogeneity of the residual variance is assumed to be due to measured covariates.

An alternative is a mixed-effects location scale model in which the residual variance is assumed to be due to unmeasured covariates, in addition to possibly being due to measured covariates, by including a random subject effect in a model for the residual variance (Hedeker et al., 2008; Hedeker & Nordgren, 2013). For example, in a model that assumes independent residuals between occasions, an exponential function (that ensures the result is positive) can be used to model the residual variance (cf. Hedeker et al., 2008):

$$\sigma_i^2 = exp\left(\tau_0 + v_i\right)$$

where the exponentiated value of τ_0 is the within-person variance for a person whose random effect v_i is equal to 0. The random effect v_i is assumed to be i.i.d. lognormal across subjects. Similar to earlier work (e.g., Harvey, 1976), the model for the residual variance can be expanded to include measured covariates. For example, let X_{ij} and W_i denote a time-varying covariate and a between-subjects covariate, respectively, a model for σ^2_{ij} that is assumed to vary by individual and according to the two covariates is

$$exp\left(\tau_0 + \tau_1 X_{ij} + \tau_2 W_i + v_i\right),$$

where τ_0, when exponentiated, is the within-person variance for a person whose random effect v_i, within-subject covariate X_{ij}, and between-subject covariate W_i are equal to 0. The coefficients τ_1 and τ_2 are the effects of the within- and between-subject covariates, respectively, on the residual variance, with a positive effect indicating that a higher level of a covariate corresponds to greater within-subject variation and a negative effect indicating that a higher level of a covariate corresponds to a lower degree of within-subject variation. For example, using a mixed-effects location scale model, Blozis et al. (2020) reported greater between-subject heterogeneity of the within-subject variance of daily time spent engaged in leisure activities on weekends versus weekdays and for women versus men. The random effect v_i is the residual from the regression and is assumed to be independent and lognormally distributed between subjects. Importantly, the variance of the random scale effect v_i, if different from zero, reflects an additional source of between-subject variation due to unobserved sources, conditional on the observed covariates.

As mentioned previously, earlier reports have documented the impact of mis-specifying the level-1 residual covariance structure of a mixed-effects model, including reports of relatively poor model fit, and perhaps more importantly, consequences for the estimated variances of the random subject effects at the second level. Ignoring correlations between residuals at the occasion level has, for example, been shown to result in overestimation of the variances of the random effects at the second level of a linear mixed-effects model (Chi & Reinsel, 1989; Ferron et al., 2002; Sivo et al., 2005). Thus, the connection between how the occasion-level covariance structure is specified and its impact on the subject-level covariance structure deserves attention when fitting a mixed-effects model to understand between-subject differences in a behavior studied over time. This point may be of particular importance when considering more advanced versions of a mixed-effects model, namely a mixed-effects location scale model. The implication is that ignoring serial correlations between the residuals at the occasion level may have consequences for the estimated variances of the random effects at the subject level, including the between-subject random scale variance that is a key component of a mixed-effects location scale model.

Between-subject covariance structure

Mixed-effects models are subject-specific models because one or more of the coefficients of a growth function are assumed to vary between subjects. The between-subject covariance matrix of the random coefficients characterizes the degree to which individuals differ in the coefficients, as well as the extent to which the coefficients covary with one another. A useful aspect of a mixed-effects model is that it is possible to test if the random coefficients are related to subject-level covariates. This could be done to test if features of a growth or developmental process are associated with individual difference measures. If subject-level covariates are included in the regression equation of a random effect, the residual of that equation is a conditional random effect, and the variance of the conditional random effect represents variation in the random effect left unaccounted for.

As stated earlier, the variance of a random coefficient may be studied as a function of occasion- and subject-level covariates (Hedeker & Nordgren, 2013). In this way, it is possible to study the determinants of the variances of the random coefficients. To illustrate this, let ϕ_0^2 denote the variance of a random intercept that is modeled as a function of an occasion- and a subject-level covariate. Using an exponential function to model the variance (cf. Hedeker & Nordgren, 2013),

$$\phi_0^2 = exp\left(\alpha_{00} + \alpha_{01}X_{ij} + \alpha_{02}W_i\right)$$

where the exponentiated value of α_{00} is the variance of the random intercept when the covariates X_{ij} and W_i are equal to 0, and α_{01} and α_{02} are the effects of X_{ij} and W_i, respectively, on the variance. A positive covariate effect would indicate greater between-subject variation in the random intercept with an increase in the covariate, and a negative effect would indicate a decrease in between-subject variation with an increase in the covariate. In a study of positive affect in adolescent cigarette smokers, for example, Hedeker et al. (2008) reported greater between-subject variation in a random intercept for individuals identified as loners relative to others and less variation among novelty seekers and 10th grade students relative to others.

Estimation

Estimation of linear mixed-effects models can be conducted using statistical software programs that use methods applicable to the estimation of other linear multivariate models, such as linear structural equation models with latent variables (Blozis, 2007). In a mixed-effects location scale model, the variances at the occasion level and the variances of the random coefficients at the subject level are typically expressed using exponential functions, and thus involve the estimation of nonlinear parameters, and more generally, estimation of a nonlinear mixed-effects model. Among the software packages developed for maximum likelihood (ML) estimation of nonlinear mixed-effects models, SAS PROC NLMIXED (Wolfinger, 1999) has been widely used to estimate mixed-effects location scale models. PROC NLMIXED is well suited for the estimation of these models because the procedure makes it convenient to include the nonlinear models for the variance of a random effect at the subject level and the variance of the residual at the occasion level to model heterogeneity of variance at both levels (Hedeker et al., 2008).

PROC NLMIXED is also adaptable for ML estimation of the mixed-effects location scale models developed here in which the residual covariance structures do not conform to the procedure's default model specifications. Specifically, we consider models in which the residual covariance structure assumes an first-order autoregressive (AR(1)) structure to help address the possibility of correlation between adjacent residuals at the first level of a model, a possible indication that the explanatory portion of a growth model does not fully account for the within-subject variation. The default residual covariance structure in PROC NLMIXED is one in which the residuals are assumed to be independent between occasions with constant variance across occasions and subjects. The GENERAL model statement option permits a user-defined likelihood function of a given model. Using this option, it is possible to specify alternative residual covariance structures, such as an AR(1) structure with a fixed variance and autocorrelation coefficient (Harring & Blozis, 2014) or a structure with a random scale and autocorrelation as discussed in this paper.

PROC NLMIXED offers ML estimation using Gaussian quadrature and a dual quasi-Newton optimization routine (SAS Institute Inc., 2015). To fit the models presented in this paper, we relied on guidance provided by Kiernan et al. (2012) by providing reasonable starting values. For each of the data sets, we started by fitting a fixed-effects model to obtain estimates of the fixed effects of the growth curve and the effects of the covariates on growth parameters. We then built up models by adding random effects, one at a time, and updating the starting values for each model based on estimates obtained for previous models. Syntax for fitting models to the flight simulation data (the second example) is in the Appendix. Tests were carried out using a type 1 error rate of .05.

In the examples that follow, SAS PROC NLMIXED was used for estimation. The default method of estimation in this procedure involves approximating the marginal loglikelihood using an adaptive Gaussian quadrature method (Pinheiro & Bates, 1995). In a given problem, a good approximation requires an adequate number of quadrature points and appropriate centering and scaling of the abscissas for the random effects. The adaptive method has the advantage of requiring fewer quadrature points and has been shown to perform well when good starting values are provided and the number of random effects is not large (Pinheiro & Bates, 1995). For some of the models considered here, estimation using adaptive Gaussian quadrature (including the Laplace approximation) stopped with a report that no valid parameter points were found. We instead used Gaussian–Hermite quadrature that approximates each integral by a weighted average of the integrand that is evaluated at specific points over a grid centered at 0. Following guidance by Carlin et al. (2001) to use a high number of quadrature points when using nonadaptive Gaussian quadrature, models were estimated using 30 quadrature points. Although an increase in the number of grid points can increase the precision of the approximation of the integral, an increase can result in a computationally intensive analysis. We observed increasing stability in the parameter estimates as the number of quadrature points was increased, but recommend caution when fitting such complex methods using this method of estimation.

Examples

The first of three data sets is analyzed using models that test possible autocorrelation and between-subject heterogeneity of the residual covariance structure at the first level of a model.

This first example is one in which covariates are not available for study; thus the example illustrates the utility of considering alternative residual covariance structures at the first level of the model. The second and third data sets include individual difference measures that are studied in relation to learning. These two examples include individual difference measures that are used to study how such measures may serve as determinants of the different aspects of variation in responses at the subject level.

Each of the three examples involves response data that tend to follow nonlinear trends. One of the appealing aspects of fitting a growth model to repeated measures is that a large number of data values per subject may be summarized using a function that is parameterized by a relatively small number of coefficients. With the addition of random coefficients, the corresponding covariance structure imposed by a mixed-effects model is a parsimonious representation of the variances and covariances of the data. It may not be reasonable, however, to assume that conditional on the subject-specific growth model that the residuals are independent. That is, although a given growth model may do well in summarizing trends in the individual-level responses, it may not be reasonable to assume that it completely addresses within-subject dependencies, especially given that a function based on a small number of parameters is used to summarize a relatively large number of data points. With a need to balance parsimony and interpretability in selecting a growth model, it may be reasonable to consider alternative covariance structures at the first level of the model with the sensible acknowledgement that the imposed structure, including AR(1), is not assumed to be the specific structure that generated the dependencies in the data, but rather, that the AR(1) process may help to better represent the correlation structure of the data.

Example 1: Performance on a complex procedural learning task

Data from a procedural learning task described in Woltz (1988) represent response latencies for 393 participants on 11 learning trial blocks (64 trials per block) (henceforth referred to as "trials") that occurred in one session. Data from the first trial are excluded from analysis because it is assumed that these scores reflect participants' adjustments to the task. Supplemental Table S1 gives descriptive statistics for trials 2–11. A plot of scores for 16 selected participants in Fig. 1 suggests that response latencies decrease at a nonconstant rate as trials progress, with scores leveling off towards the latter part of the session. We fit a series of models to these scores while making different assumptions about the residual covariance structure at the trial level, including the addition of an AR(1) structure and heterogeneity of the residual variance.

Analysis of the first data set began by fitting an exponential and a logistic growth function to the scores to test which best accounted for the data while assuming that the trial-level residuals were independent, with constant variance across trials and subjects. Each function included three subject-specific parameters representing the initial level, a lower asymptote, and a learning rate parameter. Responses were positively skewed within trials due to a subset of participants having relatively long response times. Given this, each function was applied to the data assuming that scores followed one of three response distributions— normal, gamma, or lognormal—with the latter two being continuous and positively skewed distributions. Based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC)[1] fit indices, the exponential growth function that assumed that the residuals were lognormal provided the best fit. Using log-transformed (base 10) scores (assumed to be normally distributed), this model was provisionally taken as the best fitting and used to test different assumptions about the residual covariance structure.

The exponential growth model for response y_{ij} was specified as

$$y_{ij} = \beta_{1i} - \left(\beta_{1i} - \beta_{0i}\right) exp\left\{-\beta_{2i} t_{ij}\right\} + \varepsilon_{ij}, \tag{2}$$

where t_{ij} denotes the j^{th} trial for subject i. For subject i, β_{0i} is the performance score at the first trial, β_{1i} is the potential performance level (lower asymptote), and β_{2i} is the rate parameter that combines with t_{ij} to represent the learning rate. At the subject level, each coefficient was a function of a fixed and random effect:

$$\beta_{0i} = \gamma_{00} + u_{0i}, \tag{3a}$$

$$\beta_{1i} = \gamma_{10} + u_{1i}, \tag{3b}$$

$$\beta_{2i} = \gamma_{20} + u_{2i}, \tag{3c}$$

where γ_{00}, γ_{10}, and γ_{20} are the response level at the first trial, the potential performance level, and the rate parameter, respectively, for a subject whose random effects are equal to 0.

Next, we fit two sets of models using the exponential growth function in (2). These are summarized here, followed by more detailed descriptions. Models in the first set were mixed-effects models and those in the second were mixed-effects location scale models. The first model in each set assumed that the residuals were independent between trials, and the second in each set assumed that the residuals were correlated between trials. All three growth coefficients in the first and second models of each set were random, as in (3a)–(3c). The third model in each set assumed that the

[1] Singer and Willett (2003) recommend using the level 1 sample size in calculating the BIC = −2loglikelihood + ln(N)k, where N is the number of units at level 1.

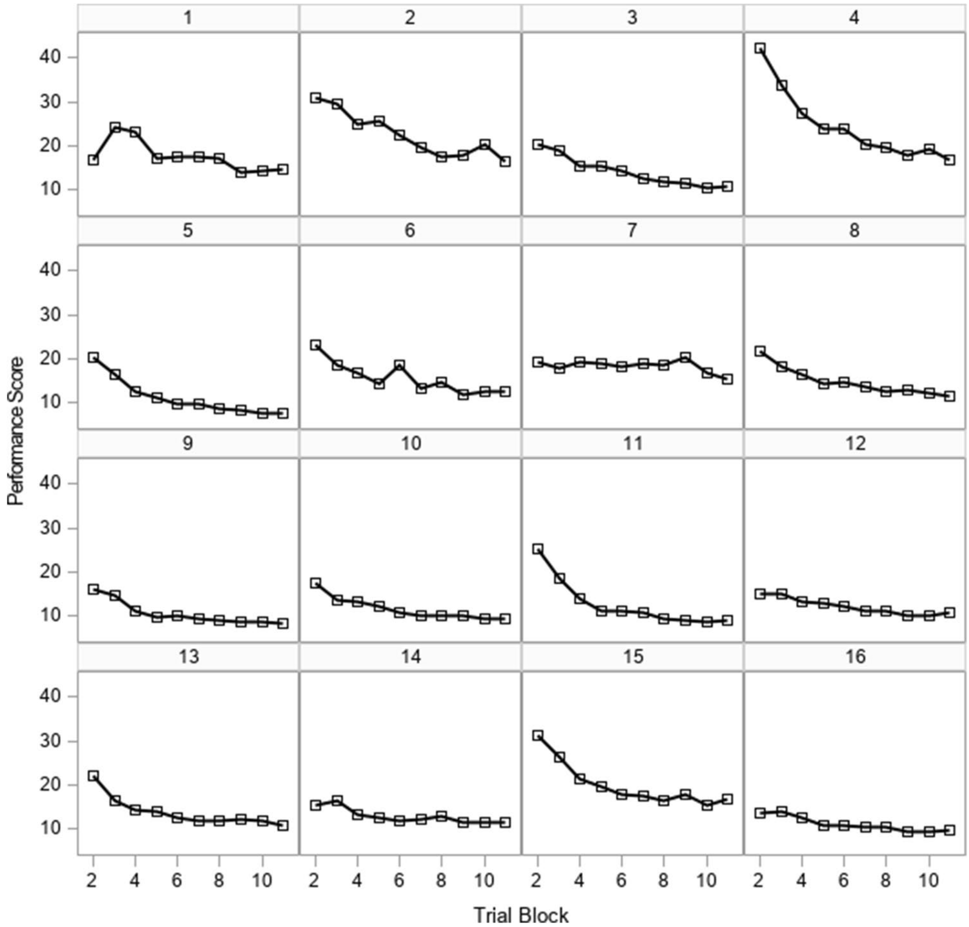

Fig. 1 Response latencies on a procedural learning task for a selection of 16 participants

residuals were correlated between trials but that the rate parameter of the growth model was fixed. By fitting these particular models, we could test whether the growth function could be simplified in terms of the number of random coefficients in exchange for using an AR(1) structure (cf. Chi & Reinsel, 1989) and whether there was evidence of heterogeneity of the residual variance across subjects. We next provide greater detail about these models.

In the first set of models, Model 1a assumed that the residuals in (2) were i.i.d. normal and independent between trials with constant variance across the 10 trials: $\boldsymbol{\Theta}_{\varepsilon} = \sigma_{\varepsilon}^2 \mathbf{I}_{10}$, where σ_{ε}^2 is the common variance. The covariance matrix of the random effects at the subject level was assumed to be homogeneous across subjects and specified as

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{u_0}^2 & & \\ \phi_{u_1 u_0} & \phi_{u_1}^2 & \\ \phi_{u_2 u_0} & \phi_{u_2 u_1} & \phi_{u_2}^2 \end{bmatrix}, \tag{4}$$

where $\phi_{u_0}^2$, $\phi_{u_1}^2$, and $\phi_{u_2}^2$ are the variances of the random intercept, asymptote, and rate parameter, respectively, and $\phi_{u_1 u_0}$, $\phi_{u_2 u_0}$, and $\phi_{u_2 u_1}$ are their covariances.

In Model 1b, the residual covariance matrix was assumed to follow an AR(1) structure:

$$\boldsymbol{\Theta}_{\varepsilon} = \sigma_{\varepsilon}^2 \begin{bmatrix} 1 & & & & \\ \rho & 1 & & & \\ \rho^2 & \rho & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \rho^9 & \rho^8 & \rho^7 & \cdots & 1 \end{bmatrix}, \tag{5}$$

where σ_{ε}^2 is a common variance assumed to be constant across trials and subjects and ρ is a fixed autocorrelation coefficient. The covariance matrix of the random effects of the growth model was assumed to have the same structure as in (4). Model 1c assumed the AR(1) residual structure in (5). The growth model assumed a random intercept and asymptote and a fixed rate parameter, thus reducing the dimensions of $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi^2_{u_0} & \\ \phi_{u_1u_0} & \phi^2_{u_1} \end{bmatrix}.$$

The next set of models were mixed-effects location scale models. In Model 2a, the residuals were assumed to be independent between trials with constant variance across trials and between-subject heterogeneity of the residual variance: $\boldsymbol{\Theta}_{\varepsilon_i} = \sigma^2_{\varepsilon_i}\mathbf{I}_{10}$, where

$$\sigma^2_{\varepsilon_i} = exp\left(\tau_0 + v_i\right). \tag{6}$$

The exponentiated value of τ_0 is the residual variance for a subject whose random effect v_i is equal to 0. The random scale effect v_i accounts for heterogeneity of variance due to unmeasured sources and is assumed to be lognormally and independently distributed between subjects. Given the random scale effect in (6), the covariance matrix of the random effects at the subject level includes the variance of the random scale effect and its covariances with the random growth coefficients:

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi^2_{u_0} & & & \\ \phi_{u_1u_0} & \phi^2_{u_1} & & \\ \phi_{u_2u_0} & \phi_{u_2u_1} & \phi^2_{u_2} & \\ \phi_{vu_0} & \phi_{vu_1} & \phi_{vu_2} & \phi^2_v \end{bmatrix}, \tag{7}$$

where ϕ^2_v is the variance of the random scale effect. Its covariances with the random growth coefficients are the first three elements in the last row of the matrix.

In Model 2b, the covariance matrix of the residuals followed an AR(1) structure for which the residual variance could vary by subject:

$$\boldsymbol{\Theta}_{\varepsilon_{ij}} = \sigma^2_{\varepsilon_{ij}} \begin{bmatrix} 1 & & & & \\ \rho & 1 & & & \\ \rho^2 & \rho & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \rho^9 & \rho^8 & \rho^7 & \cdots & 1 \end{bmatrix}, \tag{8}$$

where $\sigma^2_{\varepsilon_{ij}} = exp\left(\tau_0 + v_i\right)$. The autocorrelation coefficient was assumed to be constant across subjects. In Model 2c, an AR(1) structure as specified for Model 2b was assumed, and the rate parameter of the growth model was assumed to be fixed across subjects. The mean and covariance structures of Models 1a–1c and 2a–2c are summarized in Supplemental Table S4.

Results

Point estimates and 95% confidence intervals (CI) for Models 1a–1c and 2a–2c are in Table 1. When estimating these models, the variances at both levels were expressed by exponential functions, and so the calculated variances are provided in the lower part of the table. Models are first compared in terms of fit, and then conclusions are drawn. We first compared Model 1a that assumed independence between the residuals and Model 1b that assumed an AR(1) structure. A deviance test comparing the models was significant (χ^2(1 *df*) =148, $p < .001$), suggesting that dependencies in scores were not entirely accounted for by the growth function, and indeed, the estimated autocorrelation was .38. This result is consistent with a preference for Model 1b according to the AIC and BIC indices, where both indices are lower under Model 1b. We next compared the fit of Models 1a and 1c to test whether the number of random growth coefficients could be reduced in exchange for using an AR(1) structure (cf. Chi & Reinsel, 1989). That is, Model 1a is relatively complex due to the inclusion of three random growth coefficients, whereas Model 1c involves only two random growth coefficients and uses the AR(1) to help account for the dependencies of scores within subjects. Although the fit was better under Model 1c, suggesting that an AR(1) structure might be used to reduce the dimensionality of the model in terms of the number of random effects, the fit of Model 1b was best among the three, suggesting that the most complex model was preferred overall. Thus, assuming homogeneity of the residual variance across subjects, individuals differed in the three aspects of performance, but the subject-specific growth model did not capture all of the within-subject dependencies in scores.

Whereas Models 1a–1c are mixed-effects models that assumed homogeneity of the residual variance, Models 2a–2c are mixed-effects location scale models that permit between-subject heterogeneity of the residual variance. We compared the first model of the two sets, Models 1a and 2a, which differ in that the former assumed homogeneity of the residual variance and the latter assumed between-subject heterogeneity. The deviance test[2] was significant (χ^2(4 *df*) = 803, $p < .001$), suggesting between-subject heterogeneity of the residual variance. Deviance tests between Models 1b and 2b and between 1c and 2c (not reported here) result in comparable conclusions about the need to permit heterogeneity of the residual variance.

We next compared the fit of Models 2a and 2c to test whether the number of random growth coefficients could be reduced in exchange for using an AR(1) structure. Similar to comparisons between models in the first set, model fit according to the AIC and BIC values was better under Model 2c relative to 2a, suggesting again that an AR(1) structure might be used to reduce the dimensionality of the model in terms of the number of random effects. The fit of Model 2b, however, was the best (according to the AIC and BIC values) among the three, suggesting that

[2] Here and where other deviance tests for a nonzero variance were carried out, a 50:50 mixture of χ^2_q and χ^2_{q+1}, where q is the number of random effects at the subject level, was used to calculate the p-value (Snijders & Bosker, 2012).

Table 1 ML estimates of an exponential growth model for response latencies on a procedural learning task (n = 393)

	Mixed-effects models			Mixed-effects location scale models		
	Model 1a	Model 1b	Model 1c	Model 2a	Model 2b	Model 2c
Within-subject covariance structure	$\sigma_e^2\mathbf{I}_{10}$	$AR(1)$ with σ_e^2	$AR(1)$ with σ_e^2	$\sigma_{e_i}^2\mathbf{I}_{10}$	$AR(1)$ with $\sigma_{e_i}^2$	$AR(1)_i$ with $\sigma_{e_i}^2$
Fixed effects	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]
Initial level, γ_{00}	2.94 [2.92, 2.97]	2.93 [2.90, 2.96]	2.93 [2.90, 2.96]	2.94 [2.93, 2.95]	2.94 [2.92, 2.96]	2.94 [2.92, 2.96]
Asymptote, γ_{10}	2.30 [2.28, 2.32]	2.28 [2.25, 2.31]	2.36 [2.33, 2.38]	2.30 [2.28, 2.31]	2.31 [2.28, 2.34]	2.37 [2.35, 2.39]
Rate, γ_{20}	0.28 [0.26, 0.30]	0.24 [0.21, 0.27]	0.26 [0.24, 0.28]	0.25 [0.24, 0.26]	0.25 [0.23, 0.27]	0.28 [0.26, 0.30]
Within-subject covariance parameters						
τ_0	−10.0 [−10.1, −9.9]	−4.71 [−4.81, −4.61]	−4.59 [−4.68, −4.49]	−10.8 [−11.0, −10.6]	−5.10 [−5.22, −4.98]	−4.92 [−5.03, −4.80]
ρ		.38	.43		.34	.43
Between-subject covariance parameters						
Intercept, α_{00}	−2.56 [−2.67, −2.45]	−2.56 [−2.72, −2.41]	−2.57 [−2.72, −2.42]	−2.76 [−2.85, −2.68]	−2.61 [−2.74, −2.49]	−2.65 [−2.78, −2.52]
Asymptote, α_{10}	−3.15 [−3.25,−3.05]	−3.52 [−3.86, −3.18]	−3.17 [−3.33, −3.00]	−3.49 [−3.62, −3.37]	−3.60 [−3.84, −3.35]	−3.28 [−3.41, −3.15]
Rate, α_{20}	−4.00 [−4.18,−3.82]	−4.29 [−4.54, −4.04]		−4.28 [−4.43, −4.14]	−4.38 [−4.58, −4.19]	
Scale, α_v				0.97 [0.78, 1.15]	−0.52 [−0.71, −0.32]	−0.59 [−0.79, −0.40]
Corr(u_1, u_0), $\phi_{u_1u_0}$	.25	.60	.46	.42	.58	.52
Corr(u_2, u_0), $\phi_{u_2u_0}$	−.13	.17		.11	.11	
Corr(u_2, u_1), $\phi_{u_2u_1}$	.59	.35		.14	.06	
Corr(u_v, u_0), ϕ_{vu_0}				.14	.23	.28
Corr(u_v, u_1), ϕ_{vu_1}				.40	.48	.64
Corr(u_v, u_2), ϕ_{vu_2}				−.47	−.46	
Additional variance estimates						
Residual, $\sigma_{e_0}^2$	4.4E−05	0.01	0.01	2.1E−05	0.01	
Initial, $\phi_{u_1}^2$	0.08	0.08	0.08	0.06	0.07	0.07
Asymptote, $\phi_{u_2}^2$	0.04	0.03	0.04	0.03	0.03	0.04
Rate, $\phi_{u_3}^2$	0.02	0.01		0.01	0.01	
Scale, ϕ_v^2				2.63	0.60	0.55
−2lnL	−5792	−5940	−5871	−6595	−6727	−6641
AIC	−5771	−5917	−5855	−6565	−6696	−6618
BIC	−5732	−5874	−5824	−6511	−6638	−6575

Latency scores are log-transformed to base 10. The within-subject covariance structures are as follows: $\sigma_e^2\mathbf{I}_{10}$ denotes independence between trials and homogeneity of variance across trials and subjects; $AR(1)$ with σ_e^2 denotes a first-order autocorrelation with constant variance across trials and subjects; $\sigma_{e_i}^2\mathbf{I}_{10}$ denotes independence between trials, homogeneity of variance across trials and between-subject heterogeneity of variance; $AR(1)$ with $\sigma_{e_i}^2$ denotes a first-order autocorrelation with constant variance across trial blocks and between-subject heterogeneity of variance the more complex model was preferred. Our overall conclusion about this performance measure is that there is evidence of individual differences in the three aspects of performance, but the subject-specific growth model did not fully capture the within-subject dependencies in scores, and there was significant within-subject variation about the subject-specific performance trajectories.

For these data there were slight differences in the estimated variances of the random growth coefficients between the models that allowed for autocorrelation between the level-1 residuals versus assuming independence and whether the model assumed heterogeneity of the residual variance or not (see Table 1). More remarkable, however, was the difference in the estimated variance of the random scale effect, a value that indicates the extent to which individuals differ from each other with regard to the variance of the residuals about the subject-specific trajectories. The estimated variance of the random scale effect was reduced from 2.63 under Model 2a that assumed independence of residuals between trials to 0.60 under Model 2b that assumed an AR(1) structure, suggesting the importance of considering an AR(1) structure when drawing inferences about the degree of between-subject heterogeneity of the residual variance. This result suggests that between-subject heterogeneity of the residual variance was due in part to dependencies between the residuals after accounting for systematic growth by the subject-specific model, again illustrating the importance of accounting for autocorrelation in a growth process.

Example 2: Performance on a flight controller simulation task

The second data set is from a study of motivation, cognitive abilities, and skill acquisition (Kanfer & Ackerman, 1989). For 140 participants, the set includes repeated measures on a flight controller simulation task and a battery of motivation and cognitive ability measures. The task was designed to measure skill acquisition during a 100-minute period. The repeated measures are the number of planes brought in safely every ten minutes. Scores for the first trial are excluded from analysis, as it is assumed that participants used the first trial to adjust to the task. Scores for trials 2–10 are analyzed here. Two individual difference measures, mathematics knowledge (MK) and coding speed (CS), are used to study how each is related to performance. The MK score reflects general mathematical knowledge, including algebra and geometry. CS measures processing speed and accuracy by having subjects relate numbers in a list to information provided in a graph.

Descriptive statistics for the performance scores and individual difference measures are in Supplemental Table S2. The performance scores for 16 selected participants are displayed in Fig. 2. In Harring and Blozis (2014), this set of scores was analyzed, and among the models tested, the best fitting was one that assumed scores changed according to a negatively accelerated logistic function and the level 1 residuals followed an AR(1) structure, with the covariance structured assumed to be homogeneous across participants. Thus, from their analysis, the logistic function did not fully account for dependencies in the data, and as they showed, it was important to address this in the residual covariance structure by allowing scores between trials to correlate. This was especially important when drawing inferences about the variances of the random effects at the subject level because those estimates were impacted by the assumptions made about the residuals at the first level.

For these data, we combine a mixed-effects model that includes an AR(1) structure with features of a mixed-effects location scale model. Specifically, we use the same logistic function in Harring and Blozis (2014) and add the between-subject measures, MK and CS, to test their effects on specific aspects of learning. We relax the assumption of homogeneity of the residual covariance structure across subjects by including a random effect for the scale and autocorrelation coefficient, thus permitting the within-subject variance of the residuals and the autocorrelation coefficient to differ between subjects. Additionally, the variances of the random coefficients are studied as functions of the individual difference measures to test if the variances of the random growth coefficients are related to either of these measures.

A logistic growth function that included a random intercept β_{0i}, upper asymptote β_{1i} and rate parameter β_{2i} was applied to the performance measures:

$$y_{\mathrm{ij}} = \frac{\beta_{0i}\beta_{1i}}{\beta_{0i} + \left(\beta_{1i} - \beta_{0i}\right)exp\left\{-\beta_{2i}\left(t_{ij} - 1\right)\right\}} + \varepsilon_{\mathrm{ij}}. \tag{9}$$

For subject i, β_{0i} is the performance score at the first trial, β_{1i} is the potential performance level (upper asymptote), and β_{2i} is the rate parameter that, when combined with t_{ij}, governs the learning rate across trials. These coefficients were modeled as functions of MK and CS[3]:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}MK_i + \gamma_{02}CS_i + u_{0i}, \tag{10a}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}MK_i + \gamma_{12}CS_i + u_{1i}, \tag{10b}$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21}MK_i + \gamma_{22}CS_i + u_{2i}, \tag{10c}$$

where γ_{00}, γ_{10}, and γ_{20} are the performance levels at the first trial, the potential level, and the rate parameter for a subject with MK and CS scores equal to their respective sample means. The coefficients γ_{01}, γ_{11}, and γ_{21} are the effects of MK, holding constant the effects of CS, on the three learning coefficients, respectively. The coefficients γ_{02}, γ_{12}, and γ_{22} are

[3] MK and CS were centered about their respective sample means.

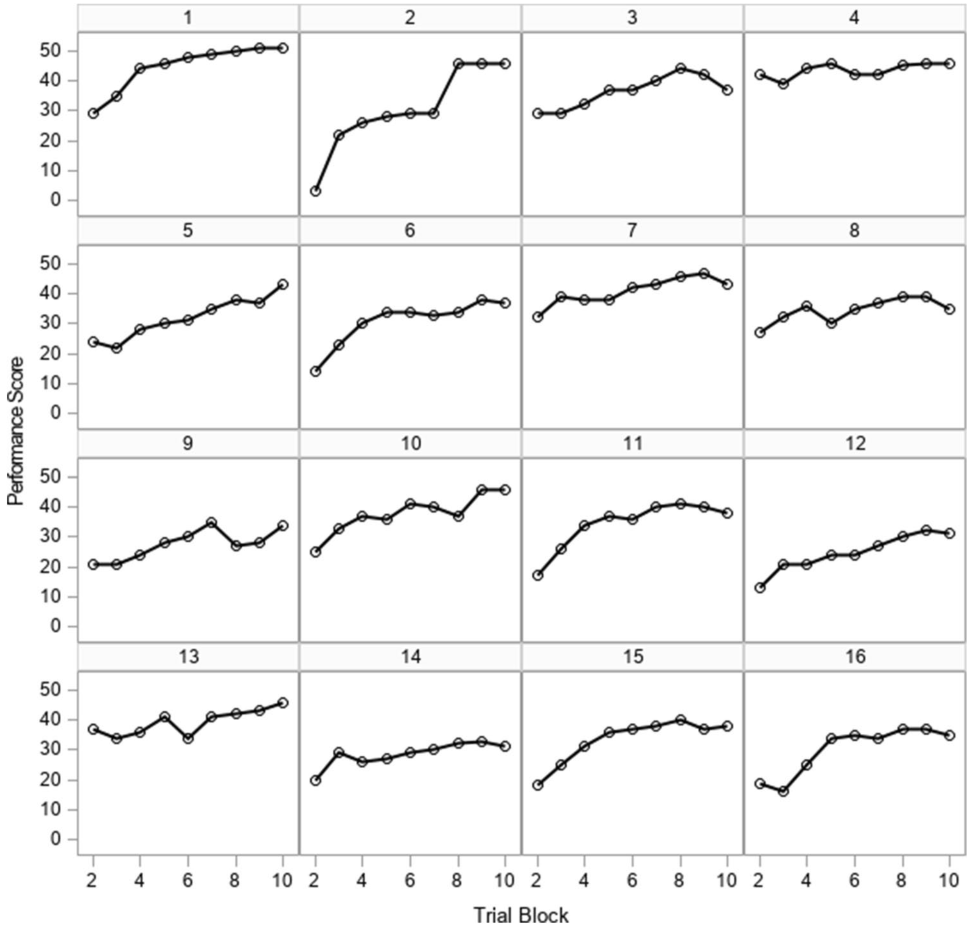

Fig. 2 Performance scores on a flight simulation task for a selection of 16 participants

the effects of CS, holding constant the effects of MK, on the three learning coefficients, respectively. The residuals of (10a)–(10c), u_{0i}, u_{1i}, and u_{2i}, are the respective subject-specific effects that remain after accounting for the effects of MK and CS.

Four models were applied to the data using the growth function in (9) with the level-2 equations in (10a)–(10c), with models differing according to the assumed covariance structure. The first two, Models 3a and 3b, were mixed-effects models that assumed homogeneity of the level 1 and level 2 covariance structures across subjects but differed in that Model 3a assumed that the residuals at the trial level were independent and Model 3b assumed that the residuals were correlated between trials (also see Harring & Blozis, 2014). The second two, Models 4a and 4b, were mixed-effects location scale models that assumed between-subject heterogeneity of the level 1 and level 2 covariance structures and differed in that Model 4a assumed that the residuals at the trial level were independent and Model 4b assumed that the residuals were correlated between trials. More details are given next.

Model 3a assumed that the residuals in (9) were i.i.d. normal and independent with constant variance across trials and subjects: $\mathbf{\Theta}_{\varepsilon} = \sigma_{\varepsilon}^2\mathbf{I}_9$, where σ_{ε}^2 is the common variance. The covariance matrix of the three conditional random effects in (10a)–(10c) was assumed to be homogeneous across subjects, similar to (4) in the first example. In Model 3b, the residual covariance matrix at level 1 was assumed to follow an AR(1) structure similar to (5), assuming homogeneity of the residual variance and autocorrelation coefficient. Similar to Model 3a, the covariance matrix of the conditional random effects in (10a)–(10c) was assumed to be homogeneous across subjects.

Model 4a was mixed-effects location scale model in which the level 1 residuals were assumed to be independent between trials with constant variance, but the variance could vary between subjects according to the measured covariates MK and CS and unmeasured sources: $\mathbf{\Theta}_{\varepsilon_{ij}} = \sigma_{\varepsilon_i}^2\mathbf{I}_9$, where

$$\sigma_{\varepsilon_i}^2 = exp\left(\tau_0 + \tau_1 MK_i + \tau_2 CS_i + v_i\right),$$

where τ_0, when exponentiated, is the residual variance for a subject whose MK and CS scores are each at their respective sample mean, and v_i is equal to 0; τ_1 and τ_2 are the effects of MK and CS, respectively, on the exponent with each adjusted for the effect of the other. The random scale effect v_i is the residual that remains after accounting for the two measured variables and was assumed to be lognormally and independently distributed between subjects. Given the added random scale effect, the level 2 covariance matrix included the variance of the random scale and its covariances with the conditional random effects of the growth model, like that shown in (7).

Model 4b was a mixed-effects location scale model that assumed correlated residuals at the trial level (similar to Model 3b) but allowed for between-subject heterogeneity of the covariance structures at levels 1 and 2. Specifically, the covariance matrix of the residuals at level 1 was assumed to follow a random AR(1) structure:

$$\mathbf{\Theta}_{\varepsilon_i} = \sigma^2_{\varepsilon_i} \begin{bmatrix} 1 & & & & \\ \rho_i & 1 & & & \\ \rho_i^2 & \rho_i & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \rho_i^8 & \rho_i^7 & \rho_i^6 & \cdots & 1 \end{bmatrix},$$

where

$$\sigma^2_{\varepsilon_i} = exp\left(\tau_0 + \tau_1 MK_i + \tau_2 CS_i + v_i\right),$$
$$\rho_i = \rho + w_i,$$

where the interpretation of the model for the within-subject variance $\sigma^2_{\varepsilon_i}$ was identical to that for Model 4a. The coefficient ρ is the autocorrelation coefficient for a subject whose random effect w_i is equal to 0; w_i is a random effect for the autocorrelation coefficient assumed to be approximated by a normal distribution. Given the random scale and random autocorrelation coefficient (v_i and w_i) of the level 1 covariance structure and the three conditional random effects of the growth model (u_{0i}, u_{1i} and u_{2i} from 10a to 10c), the covariance matrix at level 2 is as shown in (11) where, with the exception that the variance of each conditional random effect of the growth model is a function of the covariates, MK and CS:

$$\Phi = \begin{bmatrix} \phi^2_{u_0} & & & & \\ \phi_{u_1u_0} & \phi^2_{u_1} & & & \\ \phi_{u_2u_0} & \phi_{u_2u_1} & \phi^2_{u_2} & & \\ \phi_{vu_0} & \phi_{vu_1} & \phi_{vu_2} & \phi^2_v & \\ \phi_{wu_0} & \phi_{wu_1} & \phi_{wu_2} & \phi_{wv} & \phi^2_w \end{bmatrix} \quad (11)$$

$$\phi^2_{u_0} = exp\left\{\alpha_{00} + \alpha_{01}MK_i + \alpha_{02}CS_i\right\},$$
$$\phi^2_{u_1} = exp\left\{\alpha_{10} + \alpha_{11}MK_i + \alpha_{12}CS_i\right\},$$
$$\phi^2_{u_2} = exp\left\{\alpha_{20} + \alpha_{21}MK_i + \alpha_{22}CS_i\right\}.$$

The mean and covariance structures of Models 3a, 3b, 4a, and 4b are summarized in Table S1.

Results

The most complex of the four models was Model 4b. This model included a random scale effect and a random autocorrelation coefficient, thus allowing subjects to differ in the degree of variation of their residuals about their fitted trajectories and in the degree of autocorrelation between their residuals, respectively. For these data, the estimated variance of the random effect for the autocorrelation was very close to 0. From this, we concluded that subjects did not vary significantly in the autocorrelation parameter, and so Model 4b was simplified by assuming a fixed autocorrelation coefficient. Reports on Model 4b henceforth relate to this simplified model. The estimates and 95% confidence intervals from the models are given in Table 2. As variances at both levels of each model were expressed using exponential functions, Table 2 includes the calculated variances in the lower part of the table.

We first examined the overall impact of including the autocorrelation residual structure in the mixed-effects models (Models 3a and 3b) and the mixed-effects location scale models (Models 4a and 4b). Models 3a and 4a assumed independence between the residuals and Models 3b and 4b assume an AR(1) structure. A deviance test comparing Models 3a and 3b was significant[4], suggesting that dependencies in the scores were not entirely accounted for by the logistic growth function. The estimated autocorrelation under Model 3b was .39. Model 3b is also preferred to Model 3a according to the lower AIC and BIC indices under Model 3b. A significant deviance test comparing Models 4a and 4b (χ^2(1 *df*) = 27.8, $p < .001$) and lower AIC and BIC values under Model 4b suggest that dependencies in the scores were not fully accounted for by the growth function. The estimated autocorrelation under Model 4b was .44. Thus, whether one is fitting a mixed-effects model or a mixed-effects location scale model, it can be important to test the assumption of independence between residuals at the first level of a model.

Allowing for autocorrelation between the trial-level residuals impacted the estimated variances of the random effects at the subject level when comparing the mixed-effects models, as also documented elsewhere (e.g., Chi & Reinsel, 1989; Blozis & Harring, 2021). Under Models 3a and 3b, the estimated variance of the random intercept was 91.7 under Model 3a and increased to 137 under Model 3b. The variance of the random asymptote was 68.0 under Model 3a and decreased to 61.5 under Model 3b. The variance of the random rate parameter was 0.10 under both Models 3a and 3b. Despite differences in the estimated

4 (χ^2 = 16.4, p = .5*.004

Table 2 ML estimates of a logistic growth model for flight simulation performance scores (n = 140)

	Mixed-effects model		Mixed-effects location scale model	
	Model 3a	Model 3b	Model 4a	Model 4b
Within-subject covariance structure	$\sigma_e^2\mathbf{I}_9$	$AR(1)$ with σ_e^2	$\sigma_{e_i}^2\mathbf{I}_9$	$AR(1)$ with $\sigma_{e_i}^2$
Fixed effects	MLE [95% CI]	MLE[95% CI]	MLE[95% CI]	MLE[95% CI]
Initial level, γ_{10}	18.4 [16.9, 19.8]	15.3 [14.1, 16.6]	16.2 [15.8, 16.7]	17.1 [16.1, 18.0]
MK, γ_{11}	1.26 [1.16, 1.36]	1.48 [1.35, 1.60]	1.59 [1.53, 1.65]	1.86 [1.69, 2.03]
CS, γ_{12}	−0.26 [−0.28, −0.23]	−0.30 [−0.34, −0.26]	−0.22 [−0.26, −0.20]	−0.19 [−0.30, −0.09]
Asymptote, γ_{20}	39.5 [38.5, 40.5]	37.9 [36.4, 39.5]	39.1 [38.2, 39.9]	39.1 [37.6, 40.5]
MK, γ_{21}	0.59 [0.39, 0.79]	0.62 [0.29, 0.94]	0.69 [0.51, 0.87]	0.65 [0.30, 0.99]
CS, γ_{22}	0.08 [0.02, 0.15]	0.05 [−0.07, 0.17]	0.06 [0.01, 0.11]	0.09 [−0.01, 0.19]
Rate, γ_{30}	0.69 [0.62, 0.77]	0.74 [0.65, 0.82]	0.72 [0.67, 0.78]	0.69 [0.59, 0.78]
MK, γ_{31}	−0.04 [−0.04, −0.03]	−0.04 [−0.05, −0.03]	−0.05 [−0.06, −0.04]	−0.04 [−0.06, −0.03]
CS, γ_{32}	0.003 [0.002, 0.005]	0.004 [−0.002, 0.01]	0.003 [0.0002, 0.01]	0.001 [−0.005, 0.007]
Within-subject covariance parameters				
τ_0	2.22 [2.13,2.31]	2.55 [2.35, 2.74]	2.11 [1.94, 2.27]	2.54 [2.28, 2.79]
MK, τ_1			−0.04 [−0.07, −0.002]	−0.03 [−0.06, 0.005]
CS, τ_2			0.01 [−0.004, 0.02]	0.01 [−0.01, 0.02]
ρ		.39		.44
Between-subject covariance parameters				
Intercept, α_{10}	4.52 [4.37,4.67]	4.92 [4.75, 5.09]	4.50 [4.44, 4.55]	4.19 [4.05, 4.32]
MK, α_{11}			−0.08 [−0.09, −0.07]	−0.12 [−0.16, −0.08]
CS, α_{11}			0.02[0.02, 0.03]	0.03[0.01, 0.04]
Asymptote, α_{20}	4.22 [3.94,4.50]	4.12 [3.73, 4.50]	4.13 [3.94, 4.32]	3.85 [3.51, 4.20]
MK, α_{21}			−0.003 [−0.04, 0.04]	0.02 [−0.05, 0.08]
CS, α_{11}			0.01 [−0.01, 0.02]	0.02 [−0.01, 0.05]
Rate, α_{30}	−2.33 [−2.56,−2.10]	−2.27 [−2.69, −1.86]	−2.10 [−2.29, −1.91]	−2.79 [−3.36, −2.14]
MK, α_{31}			−0.15 [−0.18, −0.13]	−0.25 [−0.35, −0.15]
CS, α_{11}			−0.004 [−0.01, 0.01]	0.01 [−0.01, 0.04]
Scale, α_v			−0.67 [−1.07, −0.26]	−1.13 [−1.64, −0.62]
Corr(u_2, u_1), $\phi_{u_2u_1}$	.58	.67	.68	.59
Corr(u_3, u_1), $\phi_{u_3u_1}$	−.56	−.63	−.60	−.46
Corr(u_3, u_2), $\phi_{u_3u_2}$	−.51	−.55	−.43	−.39
Corr(v , u_1), ϕ_{vu_1}			−.32	−.31
Corr(v , u_2), ϕ_{vu_2}			−.14	−.07
Corr(v , u_3), ϕ_{vu_3}			.19	.28
Additional variance estimates				
Residual, $\sigma_{e_0}^2$	9.21	12.7	7.9	12.7
Initial, $\phi_{u_1}^2$	91.7	137.	89.7	65.7
Asymptote, $\phi_{u_2}^2$	68.0	61.5	62.3	47.1
Rate, $\phi_{u_3}^2$	0.10	0.10	0.12	0.06
Scale, ϕ_v^2			0.51	0.32
−2lnL	7346.3	7329.9	7291.8	7264.0
AIC	7382.7	7369.0	7362.5	7337.9
BIC	7425.4	7414.0	7430.2	7407.4

MK = Math Knowledge, CS = Coding Speed (each centered to their respective sample mean). The within-subject covariance structures are as follows: $\sigma_e^2\mathbf{I}_9$ denotes independence between trials and homogeneity of variance across trial blocks and subjects; $AR(1)$ with σ_e^2 denotes a first-order autocorrelation with constant variance across trial blocks and subjects; $\sigma_{e_i}^2\mathbf{I}_9$ denotes independence between trials, homogeneity of variance across trial blocks and heterogeneity of variance between subjects; $AR(1)$ with $\sigma_{e_i}^2$ denotes a first-order autocorrelation with constant variance effects of MK and CS on the coefficients of the growth model, the general interpretations are similar.

Relative to Models 3a and 3b, Models 4a and 4b were far more complex because they permitted heterogeneity of the covariance structure at both levels of the model. We first examined the estimates of the level 1 covariance structure for Models 4a and 4b. The estimated residual variance when both MK and CS were equal to their respective sample means was 7.9 under Model 4a and increased to 12.7 under Model 4b. The estimated effect of CS on the residual variance was close to 0 and not statistically significant under both Models 4a and 4b, and the effect of MK, also small in both models, was only significant under Model 4a. The most notable consequence of ignoring autocorrelation at the trial level was apparent when examining the point estimate of the variance of the random scale effect under Model 4a that assumed independence and Model 4b that assumed an AR(1) structure. The estimated variance of the random scale effect was reduced from 0.51 under Model 4a to 0.32 under Model 4b. Similar to the first example presented in this paper, this result suggests that between-subject heterogeneity of the residual variance was due in part to dependencies between the residuals after accounting for growth by the subject-specific model.

Unlike Models 3a and 3b that assumed homogeneity of the covariance structure at the subject level, Models 4a and 4b specified that the variances of the conditional random growth coefficients to be functions of the measured covariates MK and CS, and consequently, the variances of the conditional random growth coefficients are the variances when both MK and CS are at their respective sample means. Given that the effects of MK and CS are significant, it is natural to expect that the estimated variances of the conditional coefficients under Models 4a or 4b to differ from the estimates obtained under Models 3a or 3b, and indeed, the estimates do (see Table 2). Thus, we turn to study the impact of ignoring the autocorrelation in the trial-level residuals when interpreting the subject-level covariance structure.

We examined both the estimated variances of the conditional random growth coefficients and the effects of MK and CS on these variances. The estimated variance of the conditional random intercept was lower when the trial-level residuals are allowed to correlate: The estimated variance was 89.7 under Model 4a and was reduced to 65.7 under Model 4b. The estimated effects of MK and CS on this variance were also impacted by the trial-level covariance structure: The estimated effect of MK increased (in absolute value) from –0.08 under Model 4a to –0.12 under Model 4b; although slight, the estimated effect of MK increased from 0.02 under Model 4a to 0.03 under Model 4b. The degree of precision of the estimates decreased, however, when the trial-level residuals were allowed to correlate, as the estimated confidence intervals for the effects increased under Model 4b. Next, the estimated variance of the conditional asymptote was lower when the trial-level residuals were allowed to correlate: The estimated variance was 62.3 under Model 4a and 47.1 under Model 4b. Although the estimated effects of MK and CS differed between Models 4a and 4b, neither were statistically significant under either model. Finally, the estimated variance of the conditional rate parameter was lower when the trial-level residuals were allowed to correlate: The estimated variance was 0.12 under Model 4a and 0.06 under Model 4b. The estimated effects of MK and CS also differed between Models 4a and 4b: the magnitude of the effect of MK increased from –0.15 under Model 4a to –0.25 under Model 4b, and although the estimated effect of CS differed between models, the effect was not statistically significant under either model. We again note the trade-off in fitting the more complex model by the decrease in the precision of the estimates evidenced by wider estimated confidence intervals.

Example 3: Performance on a quantitative skill acquisition task

The third data set involves response latencies on a procedural learning task designed to measure quantitative skill acquisition[5]. Participants were instructed to learn a set of declarative rules for evaluating characteristics of visual stimuli that were presented in a series of trials. Each of the 12 scores represents the median time to respond across a block of 32 trials. Scores for the first trial block are not analyzed, assuming that participants were adapting to the task, leaving data for trials 2–12 for analysis. Included in a battery of individual difference measures was one of working-memory capacity obtained using a quantitative verification span paradigm. The quantitative working-memory capacity measure (QWM) is used to test how working memory capacity is related to learning acquisition.

Descriptive statistics for the response latencies and working-memory capacity are in Supplemental Table S3. Scores for 16 selected participants are displayed in Fig. 3. These scores were analyzed by applying a negatively accelerated exponential function, and like the preceding examples, alternative covariance structures are applied to understand the different sources of variation in performance scores, with working-memory capacity included in the models to test its relation to different aspects of learning and sources of score variation. We used the third data set to specifically focus on the use of an AR(1) structure as a means for reducing the dimensionality of nonlinear mixed-effects and nonlinear mixed-effects location scale models. As the dimensionality of a mixed-effects model increases, so does the computational demands. Given the relative complexity of a mixed-effects location scale model to a mixed-effects model, researchers might consider this

[5] The data were provided by Scott Chaiken of the Armstrong Laboratory, Brooks Air Force Base.

Table 2 (continued)

across trial blocks, homogeneity of the autocorrelation between subjects, and heterogeneity of variance between subjects

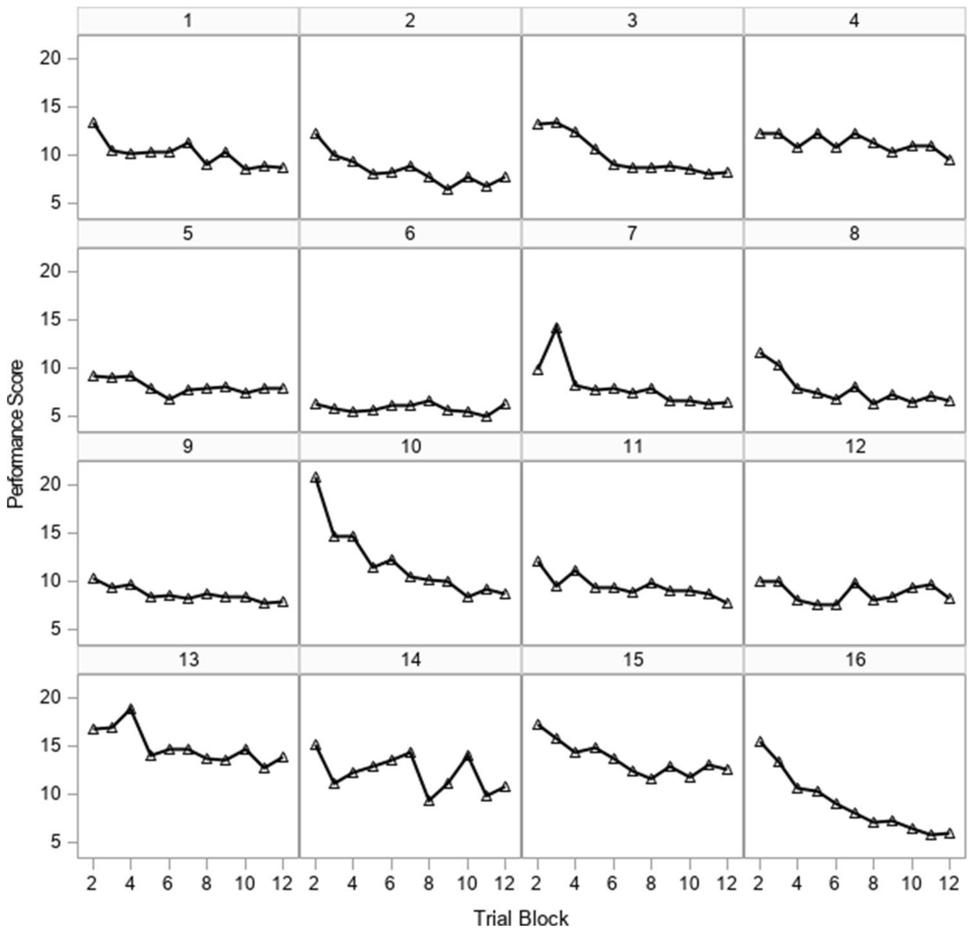

Fig. 3 Response latencies on a quantitative skill acquisition task for a selection of 16 participants

strategy to help improve the computational demands of fitting these models.

To begin, we applied three nonlinear mixed-effects models based on different growth functions to the performance scores. Among them, the exponential function[6] in (2), with each random coefficient specified as a function of QWM, provided the best fit:

$$\beta_{0i} = \gamma_{00} + \gamma_{01} QWM_i + u_{0i}, \tag{11a}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11} QWM_i + u_{1i}, \tag{11b}$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21} QWM_i + u_{2i}, \tag{11c}$$

where γ_{00}, γ_{10}, and γ_{20} are the performance levels at the first trial, potential level, and rate parameter, respectively, for a subject with QWM equal to the sample mean and whose random effects u_{0i}, u_{1i}, and u_{2i} are equal to 0. The coefficients γ_{01}, γ_{11}, and γ_{21} are the effects of QWM on the three learning coefficients. The residuals of (11a)–(11c), u_{0i}, u_{1i}, and u_{2i}, are the subject-specific effects after accounting for the effects of QWM. This model, Model 5a, assumed that the trial-level residuals were i.i.d. normal and independent with constant variance across trials and subjects: $\mathbf{\Theta}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_{11}$, where σ_ε^2 is the common variance.

Next, we fitted Model 5b that assumed a fixed rate parameter to test whether a random rate parameter was needed. A deviance test between these models suggested that the rate parameter varied between subjects ($\chi^2(3\ df) = 39, p < .001$). To then explore the idea of using an AR(1) structure to reduce the

[6] Based on AIC index values, the exponential function in Eq. (2) yielded better fit to the data overall relative to a logistic function (Eq. 7) and a Gompertz function: $y_{ij} = \beta_{1i} exp\left\{ln\left({}^{\beta_{2i}}/_{\beta_{1i}} \right) exp\left\{-t\beta_{3i}\right\}\right\} + \varepsilon_{ij}$.

number of random effects, a third model, Model 5c, was fit that assumed an AR(1) structure at level 1 and a fixed nonlinear rate parameter. We then compared the fit of Model 5a that assumed all three growth coefficients were random and that the level-1 residuals were independent between trials to the fit of Model 5c that assumed two random growth coefficients (intercept and asymptote), a fixed rate parameter, and an AR(1) structure at the trial level. According to the AIC and BIC, model fit was relatively better under Model 5c, suggesting that an AR(1) structure might be used to reduce the number of random coefficients in a growth model, and consequently the total number of model parameters and computational demands, while achieving better overall model fit. The mean and covariance structures of Models 5a–5c are summarized in Supplemental Table S4. Estimates and fit indices for Models 5a–5c are in Table 3.

Next, we followed a similar strategy in fitting models in the context of a mixed-effects location scale model. The first model, Model 6a, was a mixed-effects location scale model in which the three growth coefficients were random, as in Model 5a; the level-1 residuals were assumed to be independent with constant variance across trials, but the variance could vary by subject according to QWM and unmeasured sources: $\mathbf{\Theta}_{\varepsilon_{ij}} = \sigma^2_{\varepsilon_i}\mathbf{I}_{11}$, where

$$\sigma^2_{\varepsilon_i} = exp\left(\tau_0 + \tau_1 QWM_i + v_i\right),$$

where τ_0, when exponentiated, is the residual variance for a subject whose QWM score is equal to the sample mean and random scale effect v_i is equal to 0; τ_1 is the effect of QWM on the exponent. The random scale effect v_i is the residual after accounting for QWM and is assumed to be lognormally and independently distributed between subjects. Adding a random scale effect, the level 2 covariance matrix included the variance of the random scale and its covariances with the conditional random effects of the growth model, like (7).

A second model, Model 6b, assumed a fixed rate parameter so that we could test whether a random rate parameter was needed under this mixed-effects location scale model. A deviance test between Models 6a and 6b suggested that the rate parameter varied by subject ($\chi^2(3\ df) = 30, p < .001$). We then reduced the complexity of the growth model by fixing the rate parameter and adding an AR(1) structure to create a third model, Model 6c, where it was assumed that the autocorrelation coefficient was fixed and the residual variance was random. We then compared the fit of Model 6a that assumed all three growth coefficients were random and the level-1 residuals were independent between trials to the fit of Model 6c that assumed two random growth coefficients (intercept and asymptote), a fixed rate parameter and an AR(1) structure. According to the AIC and BIC values, model fit was better under Model 6c, suggesting that an AR(1) structure could be used to reduce the number of random coefficients in a mixed-effects location scale model while improving model fit. Estimates, 95% confidence intervals, and indices of model fit for Models 6a–6c are in Table 4. The mean and covariance structures of Models 6a–6d are summarized in Supplemental Table S4.

Discussion

The collection of repeated measures and longitudinal data is central for investigations that seek to understand change, development, or growth in measured behaviors. Mixed-effects models, a popular choice in statistical methodology, have evolved considerably since they were introduced in Laird and Ware (1982). The advantages of this major statistical framework that unify models for the population-level response and that of the individual are now numerous, including, but not limited to, the specification of linear and nonlinear models and a wide range of response distributions, handling of missing data, and data observed at different times for different subjects.

In applying a mixed-effects model to data, it is important to consider the structures necessary to address sources of heterogeneity of variance, both within and between subjects. This may be done to improve statistical inference or model fit (Blozis & Harring, 2021; Chi & Reinsel, 1989; Ferron et al., 2002; Funatogawa & Funatogawa, 2018; Harring & Blozis, 2014; Sivo et al., 2005). For instance, Chi and Reinsel show that neglecting serial correlation between level-1 residuals can result in overestimation of the variances of the random effects at the second level of a linear mixed-effects model, and Blozis and Harring showed how assumptions about the residual covariance structure (including serial correlations and heterogeneity of variance) can impact the estimated variances of the random effects at the subject level of a nonlinear mixed-effects model. Recommendations relating to statistical inference about the random-effects covariance structure at the subject level are that researchers consider alternative residual covariance structures at the first level. This is especially relevant given recent developments in mixed-effects location scale models that specifically aim to model heterogeneity of variance at both levels (Blozis et al., 2020; Hedeker & Nordgren, 2013; Williams et al., 2019).

Considering computational burden, we also considered a point of discussion in Chi and Reinsel (1989) that concerned the application of linear mixed-effects models to repeated measures data. Specifically, they discuss the use of an AR(1) structure for the level-1 residuals as a potential means to reduce the number of random effects needed to characterize a response. We explored this when fitting the nonlinear models here, including the nonlinear mixed-effects location scale models, as reducing the number of random coefficients could be helpful in reducing computational demands when fitting such complex models. Although this strategy was useful in the examples presented here, it is important to note that because one model provides a better fit than another, this is not to suggest that the best fitting model is the one that generated the data. It is simply that one might consider an alternative way

Table 3 ML estimates of a mixed-effects model for quantitative skill acquisition scores (n = 204)

	Model 5a	Model 5b	Model 5c
Within-subject covariance structure	$\sigma_e^2\mathbf{I}_{11}$	$\sigma_e^2\mathbf{I}_{11}$	$AR(1)$ with σ_e^2
Fixed effects	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]
Initial level, γ_{10}	2.46 [2.44, 2.48]	2.45 [2.42, 2.48]	2.47 [2.45, 2.50]
QWM, γ_{11}	−0.30 [−0.39, −0.21]	−0.30 [−0.46, −0.14]	−0.34 [−0.46, −0.21]
Asymptote, γ_{20}	2.00 [1.98, 2.03]	2.05 [2.02, 2.08]	2.06 [2.03, 2.09]
QWM, γ_{21}	−0.24 [−0.31, −0.17]	−0.30 [−0.40, −0.20]	−0.28 [−0.40, −0.16]
Rate, γ_{30}	0.26 [0.23, 0.30]	0.24 [0.22, 0.27]	0.25 [0.22, 0.28]
QWM, γ_{31}	−0.07 [−0.13, −0.01]	−0.19 [−0.30, −0.08]	−0.17 [−0.30, −0.05]
Within-subject covariance parameters			
τ_0	−5.0 [−5.0, −4.9]	−4.9 [−5.0, −4.8]	−4.8 [−4.9, −4.7]
ρ			.22
Between-subject covariance parameters			
Intercept, α_{10}	−2.5 [−2.6, −2.3]	−2.5 [−2.6, −2.3]	−2.4 [−2.6, −2.2]
Asymptote, α_{20}	−3.1 [−3.3 , −3.0]	−2.9 [−3.1, −2.7]	−2.9 [−3.1, −2.7]
Rate, α_{30}	−4.1 [−4.4, − 3.8]		
Corr(u_2, u_1), $\phi_{u_2u_1}$	.65	.72	.76
Corr(u_3, u_1), $\phi_{u_3u_1}$	−.16		
Corr(u_3, u_2), $\phi_{u_3u_2}$	.09		
Additional variance estimates			
Residual, $\sigma_{e_0}^2$	0.007	0.007	0.008
Initial, $\phi_{u_1}^2$	0.085	0.097	0.090
Asymptote, $\phi_{u_2}^2$	0.044	0.058	0.054
Rate, $\phi_{u_3}^2$	0.017		
−2lnL	−3421	−3382	−3430
AIC	−3395	−3362	−3408
BIC	−3352	−3329	−3372

QWM = Quantitative working-memory capacity (centered to the sample mean). The within-subject covariance structures are as follows: $\sigma_e^2\mathbf{I}_{11}$ denotes independence between trials and homogeneity of variance across trial blocks and subjects; $AR(1)$ with σ_e^2 denotes a first-order autocorrelation with constant variance across trial blocks and subjects; $\sigma_{e_i}^2\mathbf{I}_{11}$ denotes independence between trials, homogeneity of variance across trial blocks and heterogeneity of variance between subjects; $AR(1)$ with $\sigma_{e_i}^2$ denotes a first-order autocorrelation with constant variance across trial blocks, homogeneity of the autocorrelation between subjects, and heterogeneity of variance between subjects

of accounting for dependencies in the data. In the end, it is the individual researcher who makes the decision about how a model is to be specified to test particular questions about a behavior. The models presented here offer alternative methods for specifying models to test or account for heterogeneity of responses.

Estimation of mixed-effects models can be carried out using ML and Bayesian approaches. Although the current paper relies on ML, there are advantages to considering a Bayesian approach (Lin et al., 2018). For example, in one form of the model considered here, an AR(1) structure was applied to repeated measures data that permitted between-subject heterogeneity of the residual variance and the autocorrelation coefficient (see Example 2). The residual variance was modeled using an exponential function, and tests of covariate effects were carried out assuming that the effects were lognormal. Thus, this did not require special attention to the distributional assumptions made about the coefficients of the variance model. Estimation of this model using PROC NLMIXED, however, assumed that the random effect corresponding to the autocorrelation coefficient was approximated by a truncated normal distribution (with lower and upper bounds to limit the distribution between -1 and 1). Using a Bayesian approach, such as by using the SAS PROC MCMC statistical software program (Chen, 2009) would increase flexibility in the assumptions made about random effects.

Appendix

/* SAS PROC NLMIXED is a procedure that can be used to estimate nonlinear mixed-effects models. For a two-level model, the default specification for the residual covariance

Table 4 ML estimates of a mixed-effects location scale model for quantitative skill acquisition scores (n = 204)

	Model 6a	Model 6b	Model 6c
Within-subject covariance structure	$\sigma^2_{e_i}\mathbf{I}_{11}$	$\sigma^2_{e_i}\mathbf{I}_{11}$	$AR(1)$ with $\sigma^2_{e_i}$
Fixed effects	MLE [95% CI]	MLE [95% CI]	MLE [95% CI]
Initial level, γ_{10}	2.46 [2.43, 2.49]	2.48 [2.46, 2.49]	2.48 [2.46, 2.51]
QWM, γ_{11}	−0.29 [−0.38, −0.20]	−0.30 [−0.39, −0.21]	−0.29 [−0.42, −0.16]
Asymptote, γ_{20}	2.02 [1.99, 2.05]	2.08 [2.06, 2.10]	2.08 [2.05, 2.10]
QWM, γ_{21}	−0.35 [−0.43, −0.28]	−0.24 [−0.31, −0.17]	−0.24 [−0.34, −0.13]
Rate, γ_{30}	0.26 [0.22, 0.29]	0.26 [0.24, 0.29]	0.27 [0.24, 0.30]
QWM, γ_{31}	−0.20 [−0.31, −0.10]	−0.14 [−0.24, −0.05]	−0.16 [−0.27, −0.04]
Within-subject covariance parameters			
τ_0	−5.1 [−5.3, −5.0]	−5.1 [−5.2, −5.0]	−5.0 [−5.1, −4.9]
QWM, τ_1	−0.48 [−1.0, 0.06]	−0.34 [−0.84, 0.16]	−0.38 [−0.89, 0.14]
ρ			.22
Between-subject covariance parameters			
Intercept, α_{10}	−2.7 [−2.9, −2.6]	−2.7 [−2.7, −2.6]	−2.6 [−2.8, −2.5]
QWM, α_{11}	−0.67 [−1.1, −0.24]	−0.78 [−1.2, −0.37]	−0.64 [−1.2, −0.04]
Asymptote, α_{20}	−3.4 [−3.7, −3.2]	−3.0 [−3.1, −2.9]	−3.0 [−3.1, −2.9]
QWM, α_{21}	−0.02 [−0.68, 0.63]	−0.18 [−0.61, 0.25]	−0.17 [−0.83, 0.48]
Rate, α_{30}	−4.4 [−4.7, −4.2]		
QWM, α_{31}	0.71 [−0.50, 1.9]		
Scale, α_v	−0.47 [−0.63, −0.33]	−0.49 [−0.66, −0.35]	−0.48 [−0.64, −0.33]
Corr(u_2, u_1), $\phi_{u_2u_1}$	.64	.67	.69
Corr(u_3, u_1), $\phi_{u_3u_1}$	−.23		
Corr(u_3, u_2), $\phi_{u_3u_2}$	.02		
Corr(v, u_1), ϕ_{vu_1}	.36	.42	.43
Corr(v, u_2), ϕ_{vu_2}	.45	.56	.58
Corr(v, u_3), ϕ_{vu_3}	−.35		
Additional variance estimates			
Residual, $\sigma^2_{e_0}$	0.006	0.006	0.007
Initial, $\phi^2_{u_1}$	0.064	0.070	0.071
Asymptote, $\phi^2_{u_2}$	0.032	0.051	0.049
Rate, $\phi^2_{u_3}$	0.012		
Scale, ϕ^2_v	0.39	0.37	0.39
−2lnL	−3613	−3583	−3637
AIC	−3571	−3551	−3603
BIC	−3501	−3498	−3546

QWM = Quantitative working-memory capacity (centered to the sample mean). The within-subject covariance structures are as follows: $\sigma^2_e\mathbf{I}_{11}$ denotes independence between trials and homogeneity of variance across trial blocks and subjects; $AR(1)$ with σ^2_e denotes a first-order autocorrelation with constant variance across trial blocks and subjects; $\sigma^2_{e_i}\mathbf{I}_{11}$ denotes independence between trials, homogeneity of variance across trial blocks and heterogeneity of variance between subjects; $AR(1)$ with $\sigma^2_{e_i}$ denotes a first-order autocorrelation with constant variance across trial blocks, homogeneity of the autocorrelation between subjects, and heterogeneity of variance between subjects

structure at the first level is that the residuals are independent within and between level-2 units. The GENERAL model statement is used to specify alternative residual covariance structures.

Below is syntax for fitting a logistic growth model to a set of performance scores from the flight simulation task (the second example presented in the paper). The data set includes 9 repeated measures for each of 140 subjects. Two individual difference covariates, Mathematics Knowledge (MK) and Coding Speed (CS), are included in the models; each is centered about their respective sample mean. The within-subject residuals are assumed to be normally distributed.

Each script corresponds to the models reported in the manuscript for the second example. */

```
b22   -0.00202
alp00 4.1658
alp01 -0.09845
alp02 0.004995
rho10 0.6007
alp10 3.8947
alp11 0.01609
alp12 0.02062
rho20 -0.3513
rho21 -0.4187
alp20 -2.5532
alp21 -0.1446
alp22 0.007881
rhov0 -0.1769
rhov1 0.02015
rhov2 0.05191
alpv  -0.892
tau0  2.3015
tau1  -0.0297
tau2  0.004737
rho   0.3143;

covariates and a random scale effect vi;
s2e=exp(tau0 + tau1*gm_MK + tau2*gm_CS + vi);

sdv=sqrt(exp(alpv));

sd0=sqrt(exp(alp00 + alp01*gm_MK + alp02*gm_CS));
sd1=sqrt(exp(alp10 + alp11*gm_MK + alp12*gm_CS));
sd2=sqrt(exp(alp20 + alp21*gm_MK + alp22*gm_CS));

ni=9; ln2pi = 1.8378770664;
rho2 = rho * rho;
  ka =  1 / (1 - rho2);
  kb = -ka * rho;
  kc =  ka * (1 + rho2);
f0 = b00 + b01*gm_MK + b02*gm_CS + u0;
f1 = b10 + b11*gm_MK + b12*gm_CS + u1;
f2 = b20 + b21*gm_MK + b22*gm_CS + u2;
yh = (f0*f1)/(f0+((f1-f0)*exp(-f2*(x[1]))));
e[1] = y[1] - yh;

w2=0; w3=0;
do j = 2 to ni;
f0 = b00 + b01*gm_MK + b02*gm_CS + u0;
f1 = b10 + b11*gm_MK + b12*gm_CS + u1;
f2 = b20 + b21*gm_MK + b22*gm_CS + u2;
yh = (f0*f1)/(f0+((f1-f0)*exp(-f2*(x[j]))));
e[j] = y[j] - yh;
w2 = w2 + e[j]*e[j-1];
w3 = w3 + e[j]*e[j];
end;
w1 = e[1] * e[1] + e[ni] * e[ni];
w3 = w3 - e[ni]*e[ni];
quad = (ka*w1 + 2*kb*w2 + kc*w3) / s2e;
lndet = ni*log(s2e) + (ni-1) * log(1-rho2);
*loglikelihood;
Li = -0.5 * (ni * ln2pi + lndet + quad);
dm=1;
model dm ~ general(Li);
random u0 u1 u2 vi ~ normal([0,0,0,0],
[sd0*sd0,
 rho10*sd0*sd1,sd1*sd1,
 rho20*sd0*sd2,rho21*sd1*sd2,sd2*sd2,
 rhov0*sd0*sdv,rhov1*sd1*sdv,rhov2*sd2*sdv,sdv*sdv])subject=subj;
bounds -1 <= rho10 rho20 rhov0 rhov1 rhov2 rho21 <=1;
```

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13428-023-02133-1.

References

Aitkin, M. (1987). Modelling variance heterogeneity in normal regression using GLIM. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 36*(3), 332–339. https://doi.org/10.2307/2347792

Baek, E., & Ferron, J. J. M. (2020). Modeling heterogeneity of the level-1 error covariance matrix in multilevel models for single-case data. *Methodology, 16*(2), 166–185. https://doi.org/10.5964/meth.2817

Baek, E., Luo, W., & Henri, M. (2020). Issues and solutions in meta-analysis of single-case design with multiple dependent variables using multilevel modeling. *The Journal of Experimental Education*. https://doi.org/10.1080/00220973.2020.1821342

Blozis, S. A. (2007). A Newton procedure for a conditionally linear mixed-effects model. *Behavior Research Methods, 39*, 695–708.

Blozis, S. A., & Harring, J. R. (2021). Fitting nonlinear mixed-effects models with alternative residual covariance structures. *Sociological Methods & Research, 50*(2), 531–566. https://doi.org/10.1177/0049124118789718

Blozis, S. A., McTernan, M., Harring, J. R., & Zheng, Q. (2020). Two-part mixed-effects location scale models. *Behavior Research Methods, 52*(5), 1836–1847. https://doi.org/10.3758/s13428-020-01359-7

Brown, H., & Prescott, R. (2015). *Applied mixed models in medicine* (3rd ed.). Wiley.

Carlin, J. B., Wolfe, R., Brown, C. H., & Gelman, A. (2001). A case study on the choice, interpretation, and checking of multilevel models for longitudinal, binary outcomes. *Biostatistics, 2*(4), 397–416. https://doi.org/10.1093/biostatistics/2.4.397

Carroll, R. J., & Ruppert, D. (1988). *Transformations and weighting in regression*. Chapman and Hall.

Chen, F. (2009). Bayesian Modeling Using the MCMC Procedure. SAS Institute Inc. https://support.sas.com/resources/papers/proceedings09/257-2009.pdf

Chi, E. M., & Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association, 84*(406), 452–459. https://doi.org/10.1080/01621459.1989.10478790

Cook, R. D., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika, 70*(1), 1–10. https://doi.org/10.2307/2335938

Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology., 58*, 615–637. https://doi.org/10.1146/annurev.psych.58.110405.085520

Ferron, J., Dailey, R., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research, 37*(3), 379–403. https://doi.org/10.1207/S15327906MBR3703_4

Fiedler-Kelly, J., & Owen, J. S. (2014). *Introduction to population pharmacokinetic / pharmacodynamic analysis with nonlinear mixed effects models*. Wiley.

Funatogawa, I., & Funatogawa, T. (2018). *Longitudinal data analysis: Autoregressive linear mixed effects models*. Springer.

Harvey, A. C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica, 44*(3), 461–465. https://doi.org/10.2307/1913974

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(4), 657–690. https://doi.org/10.1037/0021-9010.74.4.657

Kiernan, K., Tao, J., & Gibbs, P. (2012). *Tips and strategies for mixed modeling with SAS/STAT® Procedures*. SAS Institute Inc.

Harring, J. R., & Blozis, S. A. (2014). Fitting correlated residual error structures in nonlinear mixed-effects models using SAS PROC NLMIXED. *Behavioral Research Methods, 46*(2), 372–384. https://doi.org/10.3758/s13428-013-0397-z

Hedeker, D., & Nordgren, R. (2013). MIXREGLS: A program for mixed-effects location scale analysis. *Journal of Statistical Software, 52*(12), 1–38. https://doi.org/10.18637/jss.v052.i12

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics, 64*(2), 627–634. https://doi.org/10.1111/j.1541-0420.2007.00924.x

Joo, S.-H., Ferron, J. M., Moeyaert, M., Beretvas, S. N., & Van den Noortgate, W. (2019). Approaches for specifying the level-1 error structure when synthesizing single-case data. *The Journal of Experimental Education, 87*(1), 55–74. https://doi.org/10.1080/00220973.2017.1409181

Laird, N., & Ware, J. (1982). Random-effects models for longitudinal data. *Biometrics, 38*(4), 963–974. https://doi.org/10.2307/2529876

Lin, X., Mermelstein, R. J., & Hedeker, D. (2018). A 3-level Bayesian mixed effects location scale model with an application to ecological momentary assessment data. *Statistics in Medicine, 37*(13), 2108–2119. https://doi.org/10.1002/sim.7627

McNeish, D. (2021). Specifying location-scale models for heterogeneous variances as multilevel SEMs. *Organizational Research Methods, 24*(3), 630–653. https://doi.org/10.1177/1094428120913083

Nestler, S. (2021). Modeling intraindividual variability in growth with measurement burst designs. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(1), 28–39. https://doi.org/10.1080/10705511.2020.1757455

Nestler, S. (2022). An extension of the mixed-effects growth model that considers between-person differences in the within-subject variance and the autocorrelation. *Statistics in Medicine, 41*(3), 471–482. https://doi.org/10.1002/sim.9280

Nestler, S., Geukes, K., & Back, M. D. (2018). Modeling intraindividual variability in three-level multilevel models. *Methodology, 14*(3), 95–108. https://doi.org/10.1027/1614-2241/a000150

Pinheiro, J. C., & Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics, 4*(1), 12–35. https://doi.org/10.2307/1390625

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modelling using Stata* (3rd ed.). Stata Press.

SAS Institute Inc. (2015). SAS/STAT®14.1 User's Guide. SAS Institute Inc.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University

Press. https://doi.org/10.1093/acprof:oso/9780195152968.001.0001

Sivo, S., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(2), 215–231. https://doi.org/10.1207/s15328007sem1202_2

Snijders, T. A.B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage Publishers.

Stroup, W. W., Milliken, G. A., Claassen, E. A., & Wolfinger, R. D. (2018). *SAS® for mixed models: Introduction and basic applications*. SAS Institute Inc.

Williams, D. R., Zimprich, D. R., & Rast, P. (2019). A Bayesian nonlinear mixed-effects location scale model for learning. *Behavior Research Methods, 51*(5), 1968–1986. https://doi.org/10.3758/s13428-019-01255-9

Wolfinger, R. (1999). *Fitting nonlinear mixed models with the new NLMIXED procedure (SUGI 24 Conference Proceedings, Paper 287)*. Cary, NC: SAS Institute.

Woltz, D. J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117*(3), 319–331.

Wu, L. (2010). *Mixed-effects models for complex data*. Chapman and Hall/CRC.

Open practices statement The data set and syntax for fitting models presented in the example are available for download at https://osf.io/u9hmn/?view_only=c7bf32f38ced43d386c8ffb32ab50513.