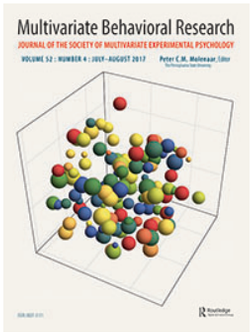

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

Explained Variance and Intraclass Correlation in a Two-Level AR(1) Model

J. Jongerling & H. Hoijtink

To cite this article: J. Jongerling & H. Hoijtink (2017) Explained Variance and Intraclass Correlation in a Two-Level AR(1) Model, Multivariate Behavioral Research, 52:4, 403-415, DOI: 10.1080/00273171.2017.1299616

To link to this article: https://doi.org/10.1080/00273171.2017.1299616

MULTIVARIATE BEHAVIORAL RESEARCH
2017, VOL. 52, NO. 4, 403–415
https://doi.org/10.1080/00273171.2017.1299616

Routledge
Taylor & Francis Group

Explained Variance and Intraclass Correlation in a Two-Level AR(1) Model

J. Jongerling[a] and H. Hoijtink[b]

[a]Department of Psychology, Education, and Child Studies, Faculty of Social Sciences, Eramus University; [b]Department of Methodology and Statistics, Faculty of Social and Behavioral Sciences, Utrecht University

ABSTRACT
The total variance of a first-order autoregressive AR(1) time series is well known in time series literature. However, despite the increased use and interest in two-level AR(1) models, an equation for the total variance of these models does not exist. This paper presents an approximation of this total variance. It will be used to compute the unexplained and explained variance at each level of the model, the proportion of explained variance, and the intraclass correlation (ICC). The use of these variances and the ICC will be illustrated using an example concerning structured diary data about the positive affect of 96 married women.

KEYWORDS
Bayesian Statistics; dynamic modeling; multilevel modeling; other topics; time series analysis

Proportion of variance explained in the AR(1) model

The AR(1) model

The AR(1) model can be used to model longitudinal data in which observations that are closer together in time are more highly correlated than observations that are further apart (Jöreskog, 1971, 1979). There are two possible specifications for an AR(1) process: a one-equation specification and a two-equation specification. If we use the one-equation specification to analyze repeated measurements of an individual that were collected on $t = 1, \ldots, T$ consecutive measurement occasions, than the model can be written as

$$y_t = c + \phi y_{t-1} + \epsilon_t \tag{1}$$

where y_t is the observed score at timepoint t, c is the intercept (i.e., the expected score when $y_{t-1} = 0$), ϕ is the AR parameter used for the regression of each observation on its immediate preceding value, and ϵ_t is the unpredictable part, referred to as the innovation, residual, or random shock. It is assumed that ϕ lies between -1 and 1 to ensure *stationarity*, that is, a situation in which the mean and variance of the process do not change over time, see Hamilton, 1994; Chatfield, 2003. If a time series is not stationary, the mean and/or variance at timepoint t can differ from the mean and/or variance at timepoint $t + k$, where k is any real number. In that case, it would not be possible to speak of *the* mean and *the* variance of the time series, as these parameters would change over time. In this study, the AR parameter ϕ is further assumed to be fixed across time ($\phi_t = \phi$ for all values of t). Finally, it is assumed that the innovations are independent and normally distributed with 0 mean and variance σ^2.

Alternatively, we can use the two-equation specification of an AR(1) model, in which y_t is viewed as consisting of two parts: a mean score μ that represents an individual's trait score (i.e., his/her long-run tendency, equilibrium, or long-term preferred state) and a error term ζ_t that represents a temporal deviation from this mean:

$$y_t = \mu + \zeta_t \tag{2}$$

The temporal deviations (or states) can subsequently be modeled with the AR(1) model:

$$\zeta_t = \phi\zeta_{t-1} + \epsilon_t \tag{3}$$

where ϕ and ϵ_t again are the AR parameter and the innovation, respectively, which are subject to the same constraints as under the one-equation specification. The temporal deviations ζ_t are assumed independent and normally distributed with mean 0 and variance σ^2_ζ.

Even though the one-equation specification and two-equation specification expressed above describe the exact same process, they are more than a simple renaming of one another. They are reparametrizations. The equivalence between these two specifications can be seen by relating the mean in Equation (2) to the intercept in Equation (1) through

$$\mu = \frac{c}{1 - \phi} \tag{4}$$

CONTACT J. Jongerling jongerling@fsw.eur.nl Department of Psychology, Education and Child Studies, Faculty of Social Sciences, Erasmus University, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands.

which is a standard result in time series literature (cf., Hamilton, 1994; Chatfield, 2003). Despite the equivalence of the two specifications, we prefer the two-equation specification since it allows for the modeling of individual means μ, which are usually more informative than intercepts c. In the rest of this article, we will therefore use the two-equation specification of the model.

Explained variance

For the AR(1) model, the total variance is

$$\sigma_y^2 = \frac{\sigma^2}{1-\phi^2} \tag{5}$$

where σ_y^2 denotes the total variance of the time series (Hamilton, 1994, pg. 53), ϕ is the AR parameter, and σ^2 is the variance of the innovation ϵ_t. Since the innovation is the unpredicted part of the AR(1) model, σ^2 can also be interpreted as the unexplained variance in an AR(1) model:

Definition 1. The unexplained variance in an AR(1) model is equal to the variance of the innovation, σ^2.

The proportion of explained variance R^2 in the AR(1) model can then be found by dividing the unexplained variance by the total variance, to get the proportion of unexplained variance, and substracting the result from 1,

$$\begin{aligned} R^2 &= 1 - \frac{\sigma^2}{\frac{\sigma^2}{1-\phi^2}} \\ &= 1 - \frac{\sigma^2(1-\phi^2)}{\sigma^2} \\ &= \phi^2 \end{aligned} \tag{6}$$

which shows that the explained variance is equal to the AR parameter squared, just like the explained variance is equal to the regression coefficient squared in standardized normal linear regression.

Total variance in the two-level AR(1) model

The two-level AR(1) model

The two-level AR(1) model can be used to model time series for $i = 1, \ldots, N$ individuals, where N denotes the total number of individuals. In this section, we begin with presenting the level 1 or *within-person part* of the two-level AR(1) model. On this level, the scores of each of the N individuals are modeled using the AR(1) model presented in the previous section. This is followed by the presentation of the level 2 or *between-persons part* of the model, in which individual differences in the level 1 parameters are modeled.

Level 1: Within-person

Like the AR(1) model presented in the previous section, the within-person part of a two-level AR(1) model can be thought of as consisting of two parts: a mean score and a temporal deviation from this mean. The only difference is that, in the two-level AR(1) model, all parameters in the equations for these two parts have a subscript i to identify the individual to which the model applies

$$y_{it} = \mu_i + \zeta_{it} \tag{7}$$

where

$$\zeta_{it} = \phi_i \zeta_{i,t-1} + \epsilon_{it} \tag{8}$$

So, in the model above, y_{it} is the observed score of individual i at timepoint t, μ_i is the mean score of individual i, ζ_{it} is the temporal deviation of individual i at timepoint t from his/her mean, ϕ_i is individual i's AR parameter, and ϵ_{it} is the innovation of individual i at timepoint t. As was the case in the single-level AR(1) model presented in the previous section, it is assumed that ϕ_i lies between -1 and 1 to ensure stationarity, that ϕ_i is fixed across time, and that the innovations are independent and normally distributed with 0 mean and variance σ_i^2. Note that we are allowing different individuals i to have their own error variance σ_i^2 (which like the AR parameter are considered fixed across time, that is, $\sigma_{it}^2 = \sigma_i^2$ for all values of t). This makes our model a heterogeneous variance model, in the sense that different individuals are allowed to have different amount of innovation variance. However, unlike standard specifications of heterogeneous variance models, these differences in variance do not have to be fully accounted for by a level 2 predictor. This makes this multilevel extension of the AR(1) model more extensive than the ones usually considered in the literature. One exception is the study by Wang, Hamaker, and Bergeman (2012), in which person-specific innovation variance was also included in the model. However, the authors did not consider the need for this in depth. More information on this specification of a multilevel AR(1) model and the reasons for allowing for individual differences in the innovation variance are given in Jongerling, Laurenceau, and Hamaker (2015).

Level 2: Between-person

When analyzing different individuals, these individuals are likely to have different values for the parameters of the within-person model, that is, the mean μ_i, the AR parameter ϕ_i, and the innovation variance σ_i^2. These individual differences in model parameters can be modeled at level

2 of our two-level AR(1) model

$$\begin{bmatrix} \mu_i \\ \phi_i \\ \sigma_i^2 \end{bmatrix} \sim MVN\left(\begin{bmatrix} \mu_\mu \\ \mu_\phi \\ \mu_{\sigma^2} \end{bmatrix}, \begin{bmatrix} \tau_\mu^2 & & \\ \tau_{\mu\phi} & \tau_\phi^2 & \\ \tau_{\mu\sigma^2} & \tau_{\phi\sigma^2} & \tau_{\sigma^2}^2 \end{bmatrix} \right) \tag{9}$$

where MVN denotes a multivariate normal distribution[1], τ_μ^2, τ_ϕ^2, and $\tau_{\sigma^2}^2$ are the interindividual variances in the model parameters, $\tau_{\mu\phi}$ is the covariance between the mean and the AR parameter, $\tau_{\mu\sigma^2}$ is the covariance between the mean and the innovation variance, and $\tau_{\phi\sigma^2}$ is the covariance between the AR parameter and the innovation variance. Furthermore, μ_μ, μ_ϕ, and μ_{σ^2} are the expected values across individuals of μ_i, ϕ_i, and σ_i^2 respectively.

We can also write the above as separate equations for μ_i, ϕ_i, and σ_i^2, such that

$$\mu_i = \mu_\mu + \varepsilon_{i\mu} \tag{10}$$

$$\phi_i = \mu_\phi + \varepsilon_{i\phi} \tag{11}$$

$$\sigma_i^2 = \mu_{\sigma^2} + \varepsilon_{i\sigma^2} \tag{12}$$

where $\varepsilon_{i\mu}$, $\varepsilon_{i\phi}$, and $\varepsilon_{i\sigma^2}$ are individual deviations from μ_μ, μ_ϕ, and μ_{σ^2}, respectively, with means equal to 0; variances τ_μ^2, τ_ϕ^2, and $\tau_{\sigma^2}^2$; and covariances $\tau_{\mu\phi}$, $\tau_{\mu\sigma^2}$, and $\tau_{\phi\sigma^2}$ (where, as before, $\tau_{\mu\phi}$ is the covariance between the mean and the AR parameter, $\tau_{\mu\sigma^2}$ is the covariance between the mean and the innovation variance, and $\tau_{\phi\sigma^2}$ is the covariance between the AR parameter and the innovation variance).

Total variance

Until now, no expression was available for the total variance of a two-level AR(1) model. An approximation for the total variances across all timepoints t and all individuals i is derived in Appendix A

$$\sigma_y^2 \approx \frac{\mu_{\sigma^2}}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4]}{(1-(\mu_\phi^2+\tau_\phi^2))^3} + \tau_\mu^2 \tag{13}$$

Note that interindividual variance in the innovation variance does not directly influence the total variance. This follows from the fact that Equation (13) does not contain the term $\tau_{\sigma^2}^2$. Instead, random variance in the innovation variances $\tau_{\sigma^2}^2$ only influences σ_y^2 through its covariance $\tau_{\phi\sigma^2}$ with the AR parameter ϕ. Further note that not all terms in Equation (13) are easy to interpret. The interpretation of τ_μ^2, the last term in Equation (13), is straightforward. This term represents the contribution of interindividual differences in mean scores to the total variance across time series. What might be less obvious is that the first three terms represent the combined, and interrelated, contribution of the AR parameter and the innovation variances to the total variance, with the second term showing how the simple covariance between the innovation variance and the AR parameter influences the total variance. For example, this term shows that for a positive AR parameter, a negative covariance between the AR parameter and the innovation variance leads to a smaller total variance. This makes sense since for positive AR parameters, a negative covariance implies that higher AR parameters are associated with smaller innovation variances. Since both small innovation variances and high positive values of the AR parameter lead to successive scores that are more alike, the total variance will be smaller in this situation. For negative values of the AR parameter, on the other hand, the second term shows that a negative covariance leads to *more* total variance. Again this makes sense, since for negative values of the AR parameter and a negative covariance, lower (more negative) AR parameters are associated with higher innovation variance. Given that both (more) negative AR parameters (which represent switches between positive and negative scores) and larger innovation variance lead to successive scores that are less alike, the total variance should be larger in this situation.

To show that the random innovation variance indeed does not influence the total variance by itself and to get a better idea of what that means, we generated two sets of time series each for 100 individuals. In the first set, we have time series for 100 individuals in which the innovation variance is fixed, so $\tau_{\sigma^2}^2 = 0$. Specifically, the parameter values in this first set of 100 time series are $\mu_\mu = 10$, $\mu_\phi = .20$, $\mu_{\sigma^2} = 3$, $\tau_\mu^2 = 2.188$, $\tau_\phi^2 = .01$, $\tau_{\sigma^2}^2 = 0$, and $\tau_{\phi\sigma^2=0}$. Three example time series out of these 100 fixed innovation time series are plotted in Figure 1(a). In the second set of time series, all parameter values are exactly equal as in the first, but now there are interindividual differences in innovation variance with $\tau_{\sigma^2}^2 = 1$. The correlation between the innovation variances and the AR parameter are still set to 0 however. Three example time series out of these 100 random innovation time series are plotted in Figure 1(b). The total variance across these two sets of time series is equal (5.07 vs. 5.13 with the difference of .06 being caused by sampling error); however, the

[1] We use a multivariate normal distribution, and so assume the innovation variance is normally distributed instead of distributed following a more common distribution (e.g., lognormal, inverse gamma, folded normal, etc.), because the derivation of the expression for the total variance of a two-level AR(1) model requires the assumption of normality for all model parameters. We do not expect this approach to cause computational problems, because innovation variances are expected to be clearly larger than zero in the data. A derivation for the total variance of the two-level AR(1) model without the normality assumption for the innovation variance is the topic of a future study.

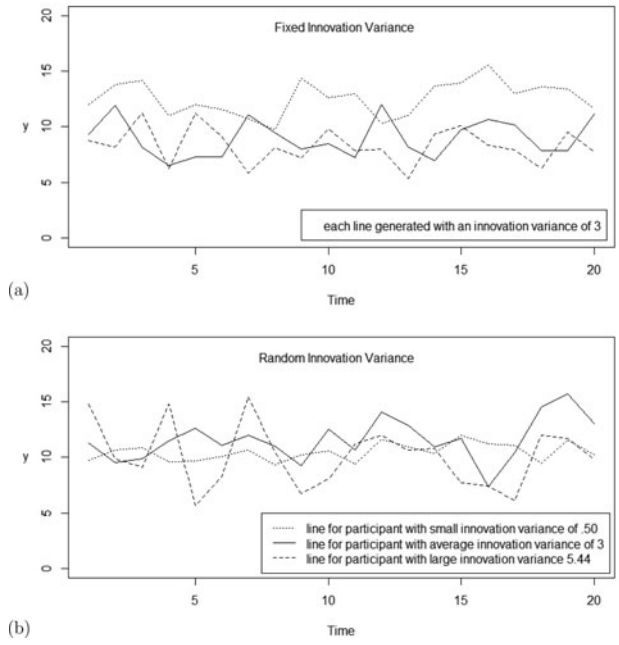

Figure 1. (a) 3 example time series out of 100 time series generated with a fixed innovation variance of 3. This results in a total variance of 5.07. (b) 3 example time series out of 100 time series generated with a random innovation variance with a mean of 3 and variance of 1. This results in a total variance of 5.13. Note that in this second set, the large innovation variances and small innovation variances cancel out to asymptotically get the same total variance as in (a).

pattern in the two sets of time series is different. To see this, letus first look at what makes up a person's score in a time series. As can be seen from Equations (7) and (8), an individual's score at timepoint t is made up of (1) his or her mean score, (2) a part that is predicted from the immediate preceding value using autocorrelation, and (3) innovation that represents random deviations from the score predicted for an individual at timepoint t by his/her mean and previous score. Both sets of time series are identical with respect to the amount of variability caused by differences in mean and differences in autocorrelation. The way that innovations influence the individual time series is different though. In the first set of time series, the size of the random deviations from the scores predicted by a person's mean and amount of autocorrelation is the same for everyone. The size of an individuals random deviations will differ from timepoint to timepoint, but the range of these deviations is the same for each of the 100 individuals in the set. In the second set, however, the range of the random deviations is not the same for everyone. Some individuals can display quite large deviations from the score predicted by his/her mean and amount of autocorrelation, while others only have small deviations from this predicted score. This is clearly visible in Figure 1.

In Figure 1(a), we see that the differences between successive timepoints in the individual time series are pretty homogeneous, and the 3 time series look quite alike. In Figure 1(b), however, the amount of difference between successive timepoints is much larger in some individuals' time series than in those of others, with the time series plotted with a solid line showing an average amount of difference between successive timepoints, the time series plotted with short dashes showing small differences between timepoints, and the time series plotted with long dashes showing large differences between successive timepoints. These last three time series also clearly look less alike than those in Figure 1(a). In short, even though there is no difference in the total variance between the time series with fixed innovation and the time series with random innovation, the total variance is less evenly distributed across individual time series when individuals differ with respect to their innovation variance. Individuals with larger ranges in innovation values contribute more variability than the average person, while individuals with smaller ranges of innovation values contribute less variability than the average persons. Both these deviations from the average cancel out, however, resulting in the same amount of total variance that is obtained for time series in which everybody has the same, average range of innovation values.

To test the validity of our approximation of the total variance of a two-level AR(1) model [Equation (13)], we undertook a small simulation study in which we determined how accurate the estimate of the total variance obtained with Equation (13) was as a measure of (1) the population variance in a two-level AR(1) model, and (2) the sample variance of the two-level AR(1) model. To do so, we generated data sets with different sample sizes and with $T = 20$ different timepoints. The parameter values chosen to generate the data will be the same for all sample sizes considered and are given in Table 1. Note that all covariances not listed in this table were set to 0. The number of timepoints and the parameter values were chosen based on literature on this kind of model from the social sciences. Specifically, the parameter values used where the ones found in the application of the two-level AR(1) model with person-specific innovation variance in Wang et al. (2012). The only exception being the correlation between the AR parameter and innovation variance $\rho_{\phi\sigma^2}$ of $-.95$, that was used in the last three simulations. This is an extreme and unrealistic value for this correlation, and it was chosen to thoroughly test our approximation. The parameter values used in this extreme case were chosen so that this extreme correlation value could be obtained.

To test Equation (13)'s ability to estimate the population variance (i.e., to test the asymptotic performance

Table 1. Results from the simulation study comparing the total variance estimated using Equation (13) to the sample variance of generated data sets with varying sample sizes and parameter estimates.

μ_μ/τ^2_μ	μ_ϕ/τ^2_ϕ	$\mu_{\sigma^2}/\tau^2_{\sigma^2}$	$\rho_{\phi\sigma^2}$	N	Sample variance	Equation (13)	Relative difference (%)
10/2.188	.20/.010	3.00/1.00	.600	20	5.2780	6.0100	13.66
10/2.188	.20/.010	3.00/1.00	.600	100	5.3111	5.4380	2.38
10/2.188	.20/.010	3.00/1.00	.600	10,000,000	5.3846	5.3788	.11
10/2.160	.20/.010	3.00/1.00	.000	20	5.2097	5.9251	13.54
10/2.160	.20/.010	3.00/1.00	.000	100	5.3454	5.4828	2.57
10/2.160	.20/.010	3.00/1.00	.000	10,000,000	5.3297	5.3242	.10
10/2.131	.20/.010	3.00/1.00	− .600	20	5.1384	5.8658	13.93
10/2.131	.20/.010	3.00/1.00	− .600	100	5.2436	5.3811	2.61
10/2.131	.20/.010	3.00/1.00	− .600	10,000,000	5.2697	5.2686	.02
10/ .220	.40/.031	2.50/1.40	− .950	20	3.1733	3.5398	11.55
10/ .220	.40/.031	2.50/1.40	− .950	100	3.2087	3.3014	2.89
10/ .220	.40/.031	2.50/1.40	− .950	10,000,000	3.2029	3.1711	1.00

Note: $\rho_{\phi\sigma^2}$ denotes the correlation between ϕ and σ^2, which is $\frac{\tau_{\phi\sigma_2}}{\tau_\phi \tau_{\sigma^2}}$. The column *Sample Variance* contains the sample variance of a generated data set. For the smaller samples of $N = 100$ and $N = 20$, we generated 100 samples for each value of $\rho_{\phi\sigma^2}$, so for these smaller samples the column *Sample variance* contains the average sample variance across the 100 data sets. The column *Equation (13)* contains the total variance estimates obtained from Equation (13). For the large samples of $N = 10,000,000$, variance estimates were obtained by entering population values (i.e., the values used to generate the data) into Equation (13) to get an estimate of the population variance. For the samples of size $N = 100$ and $N = 20$, variance estimates were obtained by entering the parameter estimates obtained in each generated data set into Equation (13) to get estimates of the sample variance. Since we generated 100 data sets for each value of $\rho_{\phi\sigma^2}$ with the smaller sample sizes, the value given in the column *Equation (13)* for these smaller samples gives the average estimate obtained from Equation (13) across the 100 data sets. The column *Relative difference* shows the difference (in percentages) between the actual variance of the samples and the variance estimate obtained from Equation (13). For the smaller samples of $N = 100$ and $N = 20$, we generated 100 samples for each value of this correlation is as follows. The sample variance between the AR parameter and the innovation variance, and the relative difference given here is the average relative difference between the sample variance and the variance estimates with Equation (13) across all samples. Data were simulated, and (where necessary) parameters were estimated using the methods described in Jongerling, Laurenceau, and Hamaker (2015).

of our approximation), we generated four very large data sets of sample size $N = 10,000,000$, each with a different value for the correlation between the AR parameter and the innovation variance. Subsequently, we determined the total variance across the 10,000,000 individuals in these samples, and compared that estimate of the variance to the one obtained by entering the parameter values used to generate the data into Equation (13). The reasoning behind this approach is that in samples of 10,000,000 individuals, the sample variance is nearly identical to the true population variance, and can therefore be used as accurate and valid benchmark for the population variance estimate obtained by entering the population parameter values (i.e., the parameter values used to generate the data) into Equation (13). If Equation (13) is correct, the total population variance estimate obtained by entering the parameter values used to generate the data into Equation (13) should be close to the large sample variance.

To test our approximation's ability to estimate sample variances in realistic situations, we generated 100 data sets with realistic sample sizes of $N = 100$ (a sample size that, like the number of timepoints, we consider realistic for these kinds of models based on literature from the social sciences [see Wang et al. (2012) and Jongerling et al. (2015)] and 100 data sets with small sample sizes of $N = 20$ for every value of the correlation between the AR parameter and the innovation variance considered. As with the large samples of $N = 10,000,000$, we subsequently compare the sample variances of these smaller samples to estimates obtained with Equation (13). Note, however, that in these smaller samples, the sample variance is a less accurate estimate of the population variance than in the large samples of $N = 10,000,000$. Comparing the sample variances from these smaller samples to estimates obtained from Equation (13) by entering the values used to generate the data therefore provides little information about the accuracy of the equation. That is why we compare the variance in these smaller samples to estimates obtained by entering sample estimates of the parameter values into Equation (13). So, with the small samples we first estimate the parameter values of our two-level AR(1) model using the data from the sample at hand, subsequently use these parameter estimates in Equation (13) to get an estimate of the total sample variance, and finally compare this estimate to the actual variance in the current sample. The differences between these two sample variance estimates are subsequently averaged across the 100 samples generated for a given value of the correlation between the AR parameter and the innovation variance, which will give an indication of how accurate Equation (13) can estimate the total sample variance under realistic conditions. The reason we generate 100 data sets of size $N = 100$ and size $N = 20$ for each value of this correlation is as follows. The sample variance between the AR parameter and the innovation variance, while we only generate one sample of $N = 10,000,000$ for each value of this correlation is that the sample variances and the performance of Equation (13) will be variable within the smaller sample sizes, while they can be considered more or less fixed in the asymptotic large sample case of

$N = 10,000,000$. In addition, we used the large sample to test our approximation's ability to estimate the population variance, while we use the smaller samples to test Equation (13)'s ability to estimate sample variances, the value of which will differ from sample to sample.

The results of this simulation study are shown in the right-hand panel of Table 1. They show that the difference between the total variances in our large samples of $N = 10.000.000$ (which can be considered accurate estimates of the true, population value of the total variance, and which are given on line 3, 6, 9, and 12 of Table 1) and the estimates of these total variances obtained by entering the parameter values used to generate the data into Equation (13), is equal to at most 1% (in case of the extreme correlation of $-.95$, where $(3.2029 - 3.1711)/3.2029 = .01$). In other words, the total variance estimate from our equation is really close to what we consider to be a good estimate of the total population variance, which indicates the validity of our approximation in Equation (13). Next, when comparing the average sample variance across the 100 smaller samples of $N = 100$ to the average total variance estimate obtained by entering the parameter estimates obtained from each of these samples into Equation (13), we see that the largest relative difference between these two is equal to 2.89% (for $\rho_{\phi\sigma^2} = -.950$). So like in the larger samples, the total variance estimate obtained from our approximation is close the sample variance in these smaller samples of $N = 100$, which proves that, next to the population variance, our approximation can accurately estimate the sample variance as well. For the small sample sizes of $N = 20$, the difference between the sample variance and the variance estimates of our approximation is larger than for the other two sample sizes. However, even with this small sample size, the relative difference is still 13.93% at most (for $\rho_{\phi\sigma^2} = -.600$). Note that it is not surprising that the relative difference is larger when comparing the sample variances from the small samples (i.e., $N = 100$ and $N = 20$) to the total variance estimate from Equation (13), than when comparing the sample variance of the large samples ($N = 10,000,000$) to the estimates obtained with our approximation. The fact that we use parameter estimates for this comparison in the smaller samples, instead of the true population values, results in an extra source of bias. Taken together, we feel these results show that Equation (13) provides accurate and useful estimates of the total sample variance.

Proportion of variance explained and intraclass correlation in the two-level AR(1) model

In the following sections, we will provide expressions for different proportions of explained variance. Specifically, we will provide expressions for (1) the proportion of explained variance at level 1, (2) the proportion of explained variance at level 2, (3) the total proportion of explained variance, and for (4) the proportion of variance explained by autocorrelation. In addition, we will provide an expression for the ICC of the two-level AR(1) model.

Proportion of explained variance at level 1

As was the case for the AR(1) model, the innovations represent the unpredicted part, and so the innovation variance can be interpreted as the unexplained variance. In contrast to the AR(1) model, however, the amount of innovation variance may differ across individuals in a two-level AR(1) model, meaning that there is an overall mean amount of innovation variance on level 1 (μ_{σ^2}), and some interindividual variability in the amount of innovation variance on level 2 ($\tau^2_{\sigma^2}$). In other words, the total innovation (or unexplained) variance is divided into two parts or parameters. One for each level of the model. We therefore define unexplained variance at level 1 of a multilevel AR(1) model as follows:

Definition 2. The unexplained variance at the first level of a two-level AR(1) model is that part of the variance of the innovations that is located at level 1. This implies

$$\sigma^2_{y,un1} = \mu_{\sigma^2} \tag{14}$$

where $\sigma^2_{y,un1}$ denotes the unexplained variance in y at level 1. Loosely formulated, each individual i (the units on level 1) has an amount of unexplained variance σ^2_i. The average of these unexplained variances is a measure of the total unexplained variance at level 1.

Following the same logic/reasoning, we define the total variance at level 1 of a two-level AR(1) model as that part of the total variance that is *not* due to between-person (i.e., level 2) variance in the model parameters:

Definition 3. The total variance at the first level of a two-level AR(1) model is that part of the total variance not due to τ^2_μ and τ^2_ϕ. Removing these level 2 variance terms from Equation (13) renders

$$\sigma^2_{y,tot1} = \frac{\mu_{\sigma^2}}{1 - \mu^2_\phi} \tag{15}$$

where $\sigma^2_{y,tot1}$ denotes the total variance in y at level 1. Loosely formulated, this expression, like Equation (1) and Equations (2) and (3), shows that a persons' score varies around his/her mean due to two different sources of fluctuation. Fluctuation due to autoregression and unexplained fluctuation. In addition, Equation (15) shows that these two sources are not simply additive.

Note that this expression is very similar to the expression for the total variance of the AR(1) model [Equation (5)]. The only difference is that in the two-level AR(1) model, the total variance at level 1 is determined using μ_ϕ and μ_{σ^2}, whereas in the AR(1) model, the total variance is determined using ϕ and σ^2. Further note that we did not include interindividual variance in the innovation variance ($\tau^2_{\sigma^2}$) in the definition of the total level 1 variance, since Equation (13) shows that this term does not affect the total variance (i.e., Equation (13) does not include the term $\tau^2_{\sigma^2}$).

The explained variance at level 1 can now be obtained by dividing the unexplained variance at level 1 by the total variation at level 1, to get the proportion of unexplained variance, and substracting the result from 1:

$$\begin{aligned} R^2_{level\,1} &= 1 - \frac{\mu_{\sigma^2}}{\frac{\mu_{\sigma^2}}{1-\mu^2_\phi}} \\ &= 1 - \frac{\mu_{\sigma^2}\left(1-\mu^2_\phi\right)}{\mu_{\sigma^2}} \\ &= \mu^2_\phi \end{aligned} \tag{16}$$

In multilevel literature, the concept of explained variance is complex (Hox, 2010, pg. 70), especially in the presence of random slopes, because maximum likelihood estimates of level 1 and level 2 variance terms might be biased (Hox, 2010, pg. 73; Snijders & Bosker, 1994). However, since this is more a computational issue than a conceptual one, we will nevertheless use this theoretical conceptualization of level 1 variance being the variance associated with level 1 parameters.

Proportion of explained variance at level 2

Definition 4. As the total variance of a two-level model is equal to the sum of the level 1 and level 2 variance (Hox, 2010, pg. 70), we define the total variance at the second level of a two-level AR(1) model as the difference between the total variance of a two-level AR(1) model [Equation (13)] and the total variance at the first level of the model [Equation (15)]:

$$\begin{aligned} \sigma^2_{y|l2} \approx &\left(\frac{1}{1-(\mu^2_\phi+\tau^2_\phi)} + \frac{(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau^4_\phi}{(1-(\mu^2_\phi+\tau^2_\phi))^3} - \frac{1}{1-\mu^2_\phi}\right)\mu_{\sigma^2} \\ &+ \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu^2_\phi+\tau^2_\phi))^2} + \tau^2_\mu \end{aligned} \tag{17}$$

Since there are no predictors at level 2 in our model, and since there are no level 1 predictors on which individuals differ in their mean scores (since $\mathbb{E}_t\left[\zeta_{i,t}\right] = \mathbb{E}_t\left[\zeta_{i,t-1}\right] = 0, \forall i$), the unexplained variance is also equal to Equation (17). Therefore,

$$R^2_{level2} = 0 \tag{18}$$

Proportion of variance related to autocorrelation

Apart from the explained variance on the first and second level, a third type of variance can be determined for the two-level AR(1) model; the variance not attributable to ϕ. We define this variance as follows:

Definition 5. The variance in a two-level AR(1) model that is not attributable to ϕ is obtained when $\mu_\phi = 0$ and $\tau^2_\phi = 0$. Applying this to Equation (13) renders

$$\sigma^2_{y|\mu\sigma^2} = \mu_{\sigma^2} + \tau^2_\mu \tag{19}$$

The proportion of the variance that is related to autocorrelation can now be determined by dividing the amount of variance that is not related to ϕ [Equation (19)] by the total variance in a two-level AR(1) model [Equation (13)] and substracting the result from 1:

$$R^2_\phi \approx 1 - \frac{\mu_{\sigma^2} + \tau^2_\mu}{\frac{\mu_{\sigma^2}}{1-(\mu^2_\phi+\tau^2_\phi)} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu^2_\phi+\tau^2_\phi))^2} + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau^4_\phi]}{(1-(\mu^2_\phi+\tau^2_\phi))^3} + \tau^2_\mu} \tag{20}$$

Total proportion of variance explained at level 1 and level 2

From Equations (16) and (18), it follows that the proportion of explained variance at level 1 is μ^2_ϕ, while the proportion of explained variance at level 2 is 0. This means that the total amount of explained variance is:

$$R^2_{total} \approx \frac{\mu^2_\phi * \sigma^2_{y,tot1}}{\frac{\mu_{\sigma^2}}{1-(\mu^2_\phi+\tau^2_\phi)} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu^2_\phi+\tau^2_\phi))^2} + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau^4_\phi]}{(1-(\mu^2_\phi+\tau^2_\phi))^3} + \tau^2_\mu} \tag{21}$$

where $*$ stands for multiplication. Since the proportion of explained variance at level 1 is equal to μ^2_ϕ, this expression is closely related to the proportion of the variance that is related to ϕ [Equation (20)]. However, the two are not the same. This is because the interindividual variance in the AR parameter τ^2_ϕ is not part of Equation (21), eventhough it is part of the variance associated with the AR parameter in Equation (20).

Intraclass correlation for a two-level AR(1) model

Using Equation (17), we can also determine the ICC for a two-level AR(1) model, that is, the percentage of the total variance of this two-level model that is located on the higher levels. The expression for the ICC is obtained

Table 2. Estimates, standard deviations, and credible intervals for the parameters of the two-level AR(1) model obtained using the data of Laurenceau, Feldman Barrett, and Rovine (2005).

	Parameter estimate	95% credibility interval
Fixed effects		
μ_μ	7.406 (.230)	6.955–7.857
μ_ϕ	.260 (.020)	.220–.300
μ_{σ^2}	4.533 (.249)	4.052–5.034
Random effects		
τ^2_μ	4.442 (.711)	3.260–6.026
τ^2_ϕ	.008 (.004)	.002–.018
$\tau^2_{\sigma^2}$	4.389 (.880)	2.948–6.391
$\tau_{\mu\phi}$	− .234 (.216)	− .625 to .216
$\tau_{\mu\sigma^2}$	.199 (.115)	− .036 to .417
$\tau_{\phi\sigma^2}$	− .057 (.246)	− .532 to .419

The table shows the parameter estimates, standard deviations, and 95% credibility intervals of the Bayesian analysis of the daily positive affect data from Laurenceau, Feldman Barrett, and Rovine (2005). The standard deviations of the posterior distributions of the parameters are given between brackets.

by dividing the total variance on level 2 [Equation (17)] by the total variance of the two-level AR(1) model [Equation (13)] to get,

$$ICC \approx \frac{\left(\frac{1}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4}{(1-(\mu_\phi^2+\tau_\phi^2))^3} - \frac{1}{1-\mu_\phi^2}\right)\mu_{\sigma^2} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} + \tau^2_\mu}{\frac{\mu_{\sigma^2}}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4]}{(1-(\mu_\phi^2+\tau_\phi^2))^3} + \tau^2_\mu}. \quad (22)$$

Empirical illustration

To illustrate the estimation of the proportions of explained variance and the ICC in the context of a two-level AR(1) model, we will analyze data collected in a study by Laurenceau, Feldman Barrett, and Rovine (2005), which were also previously analyzed in Jongerling, Laurenceau, and Hamaker (2015). In this study, spouses from 96 married couples independently completed a structured diary each evening over a period of 42 consecutive days. We summed the scores on four items labeled excited, enthusiastic, energetic, and happy (all rated on 5-point Likert scales) to comprise a single positive affect (P/A) PA score. Focusing on the women only, there were 127 out of the total of 96*42=4032 PA scores missing. Based on individual sequence plots (i.e., plots of the repeated measurements of each woman), we removed seven women who had none or very little variability over time, such that the final data set contained 89 female participants.

To analyze the data, we used a Bayesian estimation method with uninformative normal priors for μ_μ, μ_ϕ, and μ_{σ^2} and an uninformative Inverse Wishart prior for the variance covariance matrix of these three parameters. To keep the demonstration short and clear, we also assume that the (individual) AR parameters and innovation variances are fixed across time, and that the data are accurately described by an AR(1) process (while in practice, the functional form of the model and the optimal lag obviously need to be determined/tested first). A thorough explanation of this estimation method is available in Jongerling, Laurenceau, and Hamaker (2015), where it is referred to as method B1. The results are summarized in Tables 2 and 3, and in Figure 2. The first column of Table 2 contains the means and standard deviations of the posterior distribution displayed in Figure 2, while the second column contains the lower and upper bounds of the 95% central credibility intervals.

In Table 3, different types of variance and the ICC are presented. Using the parameter estimates from Table 2, it follows that the total variance is equal to 9.324, of which 4.862 is on level 1 and 4.463 is on level 2. The explained variance on level 1 is equal to 6.80%, while the total percentage of explained variance is equal to 3.50%. These amounts of explained variance are not very large, and can be seen as a strong indication that the inclusion of level 1 and level 2 predictors in the model is important (eventhough small amounts of explained variance obviously do not necessarily mean that the model is of little substantial interest. Model parameters might still be

Table 3. Variances and ICC for the data from Laurenceau, Feldman Barrett, and Rovine (2005).

	Unexplained	Explained	Total	R^2
Level 1	4.533 [Equation (14)]	.329 [Equations (15) and (16)]	4.862 [Equation (15)]	.068 [Equation (16)]
Level 2	4.463 [Equation (17)]	.000 [Equations (17) and (18)]	4.463 [Equation (17)]	.000 [Equation (18)]
Related to ϕ	.023	.329 [Equations (13) and (21)]	.349 [Equations (13) and (20)]	.037 [Equation (20)]
Total	8.996	.329 [Equations (13) and (21)]	9.324 [Equation (13)]	.035 [Equation (21)]
ICC			.479 [Equation (22)]	

Between brackets the equations used to obtain the corresponding quantities are given. Note that for the explained variances, information from two equations needs to be combined. For example, the explained variance at level 1 is calculated by first determining the total variance on this level using Equation (15), and the proportion of explained variance on this level using Equation (16). Subsequently, the specific amount of variance corresponding to this proportion of explained variance is determined by multiplying the results of Equations (15) and (16). The explained variance on level 2, the amount of explained variance related to ϕ, and the total amount of explained variance can be determined in a similar manner. The total amount of variance related to ϕ is calculated by first determining the total variance using Equation (13), and determining the proportion of variance related to ϕ using Equation (20). Subsequently, the specific amount of variance corresponding to this proportion is determined by multiplying the results of Equations (13) and (20). The amount of explained variance related to ϕ is again determined in a similar manner by multiplying the results of Equations (13) and (21). Finally, the unexplained variance related to ϕ is obtained by substracting the explained variance related to ϕ from the total variance related to ϕ.

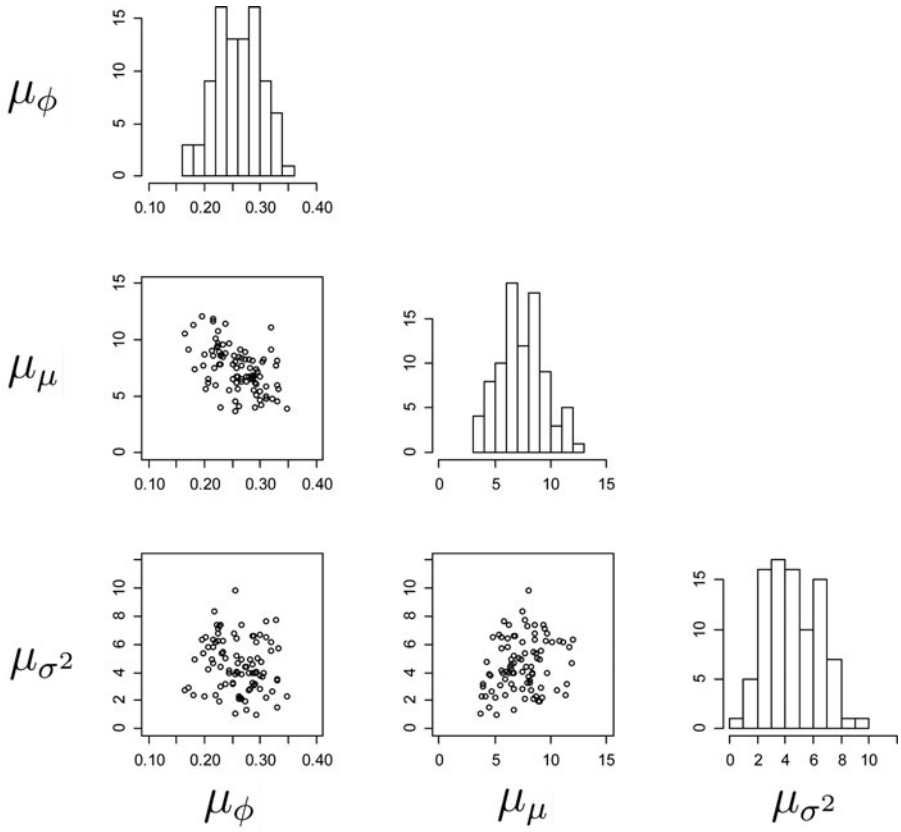

Figure 2. The histograms show the posterior distributions of μ_ϕ, μ_μ, and μ_{σ^2}. The scatterplots show the bivariate relation between the parameters of the corresponding row and column (e.g., the scatterplot on the second row of the first column shows the relation between μ_μ and μ_ϕ).

meaningfully related to important theoretical concepts). The importance of adding predictors on level 1 is further illustarted by the fact that the proportion of variance related to autoregression is equal to .037. Apparently, the only level 1 predictor currently in the model explains just 6.80% of the variance on level 1, and is related to just 3.70% of the total variance in the model. This shows that there is a lot of variance still unaccounted for. In addition, the importance of adding predictors on level 2 is further illustrated by the ICC value of .479 (indicating that 47.9% of the total variance is located on the second level). Without predictors on this level of the model, we are currently not explaining any variance at level 2, despite the fact that the amount of variance at this level is nearly half of the total variance. With such large parts of the variance on level 1 and level 2 being unrelated to the predictor in the model, it is no wonder it can explain only 3.50% of the total variance in the data.

Conclusion

In this paper, we derived an expression for the total variance of a two-level AR(1) model. This expression is an elaboration on the well-known, single-level expression for AR(1) models. Based on this expression, we derived the proportion of explained variance (both total and on level 1 and level 2 separately), the proportion of variance related to the AR-parameter ϕ, and the ICC of a two-level AR(1) model. This was demonstrated with the diary data from the positive affect study by Laurenceau, Feldman Barrett, and Rovine (2005).

Unexpectedly, the expressions derived in the study also revealed that random variance in the innovation variance does not directly influence the total variance of a two-level AR(1) model. Instead, interindividual variances in the innovation variance only influences the total variance through their correlation with the AR parameter. To us, this was an unexpected result, which has important implications. In Jongerling, Laurenceau, and Hamaker (2015), we argued that individual differences in the innovation variance τ_σ^2 are indicative of differential sensitivity and/or exposure to unmodeled factors. Taken together with the results found here, we can now conclude that such differential sensitivity and/or exposure will not show up in the total variance of an AR(1) model. Specifically, any interindividual differences in these areas that are independent of the AR parameter will go unnoticed when only looking at the total variance of the time series. This shows that the separate detection

and modeling of individual differences in innovation variances is very important, a point already made in Jongerling, Laurenceau, and Hamaker (2015) as well.

For now, we assumed that there were no (higher level) predictors in the model. Extending our two-level AR(1) model to include predictors is straightforward as it only involves writing the means of the model parameters as functions of these predictors. However, doing so would make the variance structure of the model more complicated. Expressions for the proportion of explained variance and ICC for two-level AR(1) models that include predictors are therefore the topic of future research.

We also assumed that the random innovation variance was normally distributed instead of distributed following a more common distribution like a lognormal, inverse gamma, exponential, or folded normal distribution, for example, [see Hedeker, Mermelstein, and Demirtas (2008), Hedeker, Demirtas, and Mermelstein (2009), Wang and Grimm (2012), and Estabrook, Grimm, and Bowles (2012)]. This was done because the derivation of the expression for the total variance of a two-level AR(1) model requires the assumption of normality for all model parameters. We do not expect this approach to cause computational problems because innovation variances are expected to be clearly larger than zero in the data; however, a derivation for the total variance of the two-level AR(1) model without the normality assumption would still be very useful and is a topic we are currently working on.

Finally, in our derivation and applied example, we assumed to have balanced data, with every individual having the same number of observations. Question remains if our approximation also works well with unbalanced data. We think it does because the expression for the total variance that we expand using a Taylor series [Equation (5)] applies to any time series (regardless of its length) and because the multilevel parameter estimates we use in our expression for the total variance [Equation (13)] are corrected for unbalancedness in the data. Nevertheless, testing our approximation robustness to unbalanced data is an important step, and will be part of our future research.

Article information

Conflict of Interest Disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors. institutions is not intended and should not be inferred.

References

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, *18*, 265–271.

Bohrnstedt, G., & Goldberger, A. (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association*, *64*, 1439–1442.

Chatfield, C. (2003). *The analysis of time series: An introduction* (6th ed.). Washington, DC: Chapman & Hall/CRC.

Estabrook, R., Grimm, K. J., & Bowles, R. P. (2012). A Monte Carlo simulation study of the reliability of intraindividual variability. *Psychology*, *27*, 560–576.

Hamilton, J. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Hedeker, D., Demirtas, H., & Mermelstein, R. (2009). A mixed ordinal location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Statistics and Its Interface*, *2*(4), 391–401.

Hedeker, D., Mermelstein, R., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Biometrics*, *64*, 627–634.

Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.); G. Marcoulides, Ed. New York, NY: Routledge.

Jongerling, J., Laurenceau, J. P., & Hamaker, E. L. (2015). A multilevel AR(1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research*, *50*(3), 334–349.

Jöreskog, K. G. (1971). Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology*, *23*, 121–145.

Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal-developmental investigations. In J. Nesselroade & P. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303–352). New York: Academic Press.

Laurenceau, J. P., Feldman Barrett, L., & Rovine, M. J. (2005). The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *Journal of Family Psychology*, *19*, 314–323.

Mood, A. M., Graybill, F. A., & Boes, D. C. (1985). *Introduction to the theory of statistics* (3rd ed.). London: McGraw-Hill.

Snijders, T., & Bosker, R. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, 342–363.

Wang, L. P., & Grimm, K. J. (2012). Investigating reliabilities of intraindividual variability indicators. *Multivariate Behavioral Research*, *47*, 771–802.

Wang, L. P., Hamaker, E. L., & Bergeman, C. S. (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods*, *17*, 567–581.

Appendix A

For a simple univariate AR(1) model, the well-known expression for the total variance can be written as follows:

$$\sigma_y^2 = \frac{\sigma^2}{1-\phi^2} \tag{A1}$$

where σ_y^2 is the total variance of the time series, ϕ is the AR parameter used to regress the current state on the previous one(s), and σ^2 is the error variance, referred to as the innovation variance in time series literature, which represent variance that could not be predicted based on previous scores or states (and as such can be thought of as the collection of all unmodeled factors that influence the process under investigation (Jongerling, Laurenceau, and Hamaker, 2015).

Although the expression for the total variance of a single-level AR(1) model is a well-known result in time series literature, no such expression exists for multilevel extensions of the model. To find such an expression for a two-level AR(1) model, we start with rewriting the total variance as the difference between two expected values (i.e., the expected value over time (t) of the squared repeated observations ($\mathbb{E}_t(y_t^2)$) minus the squared expectation over time of the repeated observations ($\mathbb{E}_t(y_t)^2$))[Equation (A2)]:

$$\begin{aligned}\sigma_y^2 &= \mathbb{E}_t(y_t - \bar{y})^2 \\ &= \mathbb{E}_t(y_t^2) - \mathbb{E}_t(y_t)^2 \\ &\quad \text{and so, from Equation (A1)} \\ \mathbb{E}_t(y_t^2) - \mathbb{E}_t(y_t)^2 &= \frac{\sigma^2}{1-\phi^2}\end{aligned} \tag{A2}$$

Next, we take the expected value over all individuals on both sides of the equality sign [Equation (A3)]:

$$\mathbb{E}_i[\mathbb{E}_t(y_{it}^2) - \mathbb{E}_t(y_{it})^2] = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right] \tag{A3}$$

$$\mathbb{E}_i[\mathbb{E}_t(y_{it}^2) - \mu_i^2] = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right]$$

$$\mathbb{E}_{it}(y_{it}^2) - \mathbb{E}_i(\mu_i^2) = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right] \tag{A4}$$

Since, for any random variable x

$$\mathbb{E}(x-\mu_x)^2 = \mathbb{E}(x^2) - \mathbb{E}(x)^2 \tag{A5}$$

where $\mathbb{E}(x-\mu_x)^2$ is the variance of variable x, we can rewrite the left side of Equation (A4) to get,

$$\mathbb{E}_{it}(y_{it}^2) - [\mathbb{E}_i(\mu_i)^2 + \tau_\mu^2] = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right] \tag{A6}$$

$$\mathbb{E}_{it}(y_{it})^2 + \sigma_y^2 - \mathbb{E}_i(\mu_i)^2 - \tau_\mu^2 = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right]$$

$$\sigma_y^2 = \mathbb{E}_i\left[\frac{\sigma_i^2}{1-\phi_i^2}\right] + \tau_\mu^2$$

where σ_y^2 is again the total variance of the time series, σ_i^2 is the innovation or error variance of individual i, ϕ_i is the AR parameter of individual i, and τ_μ^2 is interpersonal variation in μ. Since there are no simple exact formulas for the mean of a quotient, we eliminate the expected value of the quotient on the right side of the equation by using Taylor series (Mood, Graybill, & Boes, 1985, pg. 181). A Taylor series can be used to approximate the value of a function around a specific value, and, in general, the Taylor series for function *f(x)* around the value *a* can be written as

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{A7}$$

where $f^{(n)}(a)$ denotes the nth derivative of f evaluated at the value a. The more derivatives of *f(x)* used in the series, the more precise the approximation of *f(a)*. Here, second-order Taylor series (i.e., Taylor series in which the highest used derivative is the second-order derivative) was used. The Taylor series expansion for the product of two random variables is equal to the product of the Taylor series that can be derived for the variables separately. So, to get the Taylor series for the quotient $\frac{\sigma_i^2}{1-\phi_i^2}$ about $(\mu_{\sigma^2}, \mu_{[1-\phi^2]})$, we first consider σ_i^2 and $1-\phi_i^2$ as our two random variables, and construct separate second-order Taylor series for the function $f(\sigma_i^2) = \sigma_i^2$ and the function $g(1-\phi_i^2) = \frac{1}{1-\phi_i^2}$. For $f(\sigma_i^2)$, the second-order Taylor series around μ_{σ^2} is equal to

$$f(\mu_{\sigma^2}) = \sigma_i^2 + (\sigma_i^2 - \mu_{\sigma^2}) \tag{A8}$$

Note that the second- and higher order terms in this Taylor series are 0 because we take σ_i^2 as our parameter of interest. The first derivative of σ_i^2 is equal to 1, while the second- and higher-order derivatives are 0.

For $g(1-\phi_i^2)$, the second-order Taylor series around $\mu_{[1-\phi^2]}$ is equal to

$$\begin{aligned}g(\mu_{[1-\phi^2]}) \approx\ & \frac{1}{(1-\phi_i^2)} - \frac{1}{(1-\phi_i^2)^2}((1-\phi_i^2) - \mu_{[1-\phi^2]}) \\ & + \frac{1}{(1-\phi_i^2)^3}((1-\phi_i^2) - \mu_{[1-\phi^2]})^2\end{aligned} \tag{A9}$$

By subsequently multiplying these two Taylor series term by term, we get the following second-order Taylor series

for the quotient of σ_i^2 and $1-\phi_i^2$

$$
\begin{aligned}
&f(\mu_{\sigma^2})g(\mu_{[1-\phi^2]}) \\
&\quad\approx \frac{\sigma_i^2}{(1-\phi_i^2)} - \frac{\sigma_i^2}{(1-\phi_i^2)^2}((1-\phi_i^2)-\mu_{[1-\phi^2]}) \\
&\qquad + \frac{\sigma_i^2}{(1-\phi_i^2)^3}((1-\phi_i^2)-\mu_{[1-\phi^2]})^2 \\
&\qquad + \frac{(\sigma_i^2-\mu_{\sigma^2})}{(1-\phi_i^2)} - \frac{(\sigma_i^2-\mu_{\sigma^2})}{(1-\phi_i^2)^2}((1-\phi_i^2)-\mu_{[1-\phi^2]}) \\
&\qquad + \frac{(\sigma_i^2-\mu_{\sigma^2})}{(1-\phi_i^2)^3}((1-\phi_i^2)-\mu_{[1-\phi^2]})^2
\end{aligned}
\tag{A10}
$$

When we now take the expectation of both sides of Equation (A10), this simplifies to

$$
\begin{aligned}
\mathbb{E}_i[f(\mu_{\sigma^2})g(\mu_{[1-\phi^2]})] &\approx \frac{\mu_{\sigma^2}}{\mu_{[1-\phi^2]}} \\
&- \frac{\tau_{(\sigma^2,1-\phi^2)}}{(\mu_{[1-\phi^2]})^2} + \frac{\mu_{\sigma^2}\tau^2_{[1-\phi^2]}}{(\mu_{[1-\phi^2]})^3}
\end{aligned}
\tag{A11}
$$

Substituting this expression for the expectation of the quotient in Equation (A6) we get

$$
\begin{aligned}
\sigma_y^2 &\approx \frac{\mu_{\sigma^2}}{\mu_{[1-\phi^2]}} - \frac{\tau_{(\sigma^2,1-\phi^2)}}{(\mu_{[1-\phi^2]})^2} + \frac{\mu_{\sigma^2}\tau^2_{[1-\phi^2]}}{(\mu_{[1-\phi^2]})^3} + \tau_\mu^2 \\
&\approx \frac{\mu_{\sigma^2}}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{\rho_{(\sigma^2,\phi^2)}\tau_{\sigma^2}\tau_{\phi^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} \\
&\quad + \frac{\mu_{\sigma^2}\tau^2_{\phi^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^3} + \tau_\mu^2
\end{aligned}
\tag{A12}
$$

This takes care of all the expected values, but the expression still contains some random variables that are not directly modeled, like $\tau^2_{\phi^2}$ instead of τ^2_ϕ for example. To make sure our expression for the total variance of a two-level AR(1) model only contains parameters that are directly modeled, we view the square of the AR parameters as the product of two random variables, that is, as the product of the AR parameters with themselves. Aroian (Aroian, 1947) showed that the standard deviation of the product of two random variables θ_1 and θ_2 can be written as

$$
\begin{aligned}
&\tau_{\theta_1\theta_2} = \\
&\tau_{\theta_1}\tau_{\theta_2}\sqrt{\left(\frac{\mu_{\theta_1}}{\tau_{\theta_1}}\right)^2 + \left(\frac{\mu_{\theta_2}}{\tau_{\theta_2}}\right)^2 + 2\rho\left(\frac{\mu_{\theta_1}}{\tau_{\theta_1}}\right)\left(\frac{\mu_{\theta_2}}{\tau_{\theta_2}}\right) + 1 + \rho^2}
\end{aligned}
\tag{A13}
$$

where ρ is the correlation between the two random variables. Applying this to our 'product' of AR parameters, this expression becomes

$$
\begin{aligned}
\tau_{\phi\phi} &= \tau_\phi^2\sqrt{\left(\frac{\mu_\phi}{\tau_\phi}\right)^2 + \left(\frac{\mu_\phi}{\tau_\phi}\right)^2 + 2\left(\frac{\mu_\phi}{\tau_\phi}\right)\left(\frac{\mu_\phi}{\tau_\phi}\right) + 2} \\
&= \tau_\phi^2\sqrt{4\left(\frac{\mu_\phi}{\tau_\phi}\right)^2 + 2}
\end{aligned}
\tag{A14}
$$

Substituting this expression for τ_{ϕ^2} (and the square of this expression for $\tau^2_{\phi^2}$) in Equation (A12) results in

$$
\begin{aligned}
\sigma_y^2 &\approx \frac{\mu_{\sigma^2}}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{\rho_{(\sigma^2,\phi^2)}\tau_{\sigma^2}\sqrt{(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} \\
&\quad + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4]}{(1-(\mu_\phi^2+\tau_\phi^2))^3} + \tau_\mu^2
\end{aligned}
\tag{A15}
$$

Finally, the correlation between the innovation variance and the square of the AR parameter ($\rho_{(\sigma^2,\phi^2)}$) can also be interpreted as the correlation between a random variable (the innovation variance) and the product of a random variable (the AR parameter multiplied by itself). According to Bohrnstedt and Goldberger (1969), the covariance between the product of the two random variables x and y, and a third random variable v can be written as

$$
\tau_{xy,v} = \mathbb{E}(x)\tau_{y,v} + \mathbb{E}(y)\tau_{x,v} + \mathbb{E}\left[(\Delta x)(\Delta y)(\Delta v)\right]
\tag{A16}
$$

where $\Delta x = x - \mathbb{E}(x)$, $\Delta y = y - \mathbb{E}(y)$, and $\Delta v = v - \mathbb{E}(v)$. Since the product in our correlation is the product of one random variable with itself, and since $\tau_x^2 = (x-\mathbb{E}(x))^2$, the correlation between the innovation variance and the square of the AR parameter can be written as

$$
\rho_{(\sigma^2,\phi^2)} = \frac{2\mu_\phi\tau_{\phi\sigma^2} + \mathbb{E}_i(\tau_\phi^2(\sigma_i^2-\mu_{\sigma^2}))}{\tau_{\sigma^2}\sqrt{(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4}}
\tag{A17}
$$

where the term under the square root in the denominator can again be recognized from Equation (A14) as the standard deviation of ϕ^2, and the term $\mathbb{E}_i(\tau_\phi^2(\sigma_i^2-\mu_{\sigma^2}))$ will always be relatively small (since μ_{σ^2} is the mean value of σ^2) and can therefore be ignored.

Substitution in Equation (A15) results in the following expression for the total variance of a two-level AR(1) model

$$
\begin{aligned}
\sigma_y^2 &\approx \frac{\mu_{\sigma^2}}{1-(\mu_\phi^2+\tau_\phi^2)} + \frac{2\mu_\phi\tau_{\phi\sigma^2}}{(1-(\mu_\phi^2+\tau_\phi^2))^2} \\
&\quad + \frac{\mu_{\sigma^2}[(4(\frac{\mu_\phi}{\tau_\phi})^2+2)\tau_\phi^4]}{(1-(\mu_\phi^2+\tau_\phi^2))^3} + \tau_\mu^2
\end{aligned}
\tag{A18}
$$

Note that interindividual variance in the innovation variance does not directly influence the total variance. This follows from the fact that Equation (A18) does not contain the term $\tau^2_{\sigma^2}$. Instead, random variance in the innovation variances only influences σ^2_y through its correlation with the AR parameter as can be seen by the covariance term $\tau_{\phi\sigma^2}$ included on the right-hand side of Equation (A18).