Psicothema

Psicothema 2020, Vol. 32, No. 3, 399-409
doi: 10.7334/psicothema2020.63

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Sensitivity of five information criteria to discriminate covariance structures with missing data in repeated measures designs

Pablo Livacic-Rojas[1], Paula Fernández[2], Guillermo Vallejo[2], Ellián Tuero-Herrero[2], and Feliciano Ordóñez[2]
[1] University of Santiago de Chile and [2] University of Oviedo

Abstract

Backgrounds: This study analyzes the effectiveness of different information criteria for the selection of covariance structures, extending it to different missing data mechanisms, the maintenance and adjustment of the mean structures, and matrices. **Method:** The Monte Carlo method was used with 1,000 simulations, SAS 9.4 statistical software and a partially repeated measures design (p=2; q=5). The following variables were manipulated: a) the complexity of the model; b) sample size; c) matching of covariance matrices and sample size; d) dispersion matrices; e) the type of distribution of the variable; f) the non-response mechanism. **Results:** The results show that all information criteria worked well in Scenario 1 for normal and non-normal distributions with heterogeneity of variance. However, in Scenarios 2 and 3, all were accurate with the ARH matrix, whereas AIC, AICCR and HQICR worked better with TOEP and UN. When the distribution was not normal, AIC and AICCR were only accurate in Scenario 3, more heterogeneous and unstructured matrices, with complete cases, MAR and MCAR. **Conclusions:** In order to correctly select the matrix it is advisible to analyze the heterogeneity, sample size and distribution of the data.

Keywords: Information criteria, covariance structures, missing data, sensitivity, repeated measures designs.

Resumen

Sensibilidad de cinco criterios de información para discriminar estructuras de covarianza bajo pérdida de datos en diseños de medidas repetidas. **Antecedentes:** el presente trabajo analiza la efectividad de distintos criterios de información para seleccionar estructuras de covarianza extendiéndolo a diferentes mecanismos de pérdida de datos, la mantención y ajustes de las estructuras de medias y las matrices. **Método:** se utilizó el método Monte Carlo con 1.000 simulaciones, el software estadístico SAS 9.4 y un diseño de medidas parcialmente repetidas (p=2; q=5). Las variables manipuladas fueron: a) complejidad del modelo; b) tamaño muestral; c) emparejamiento de las matrices de covarianza y tamaño muestral; d) matrices de dispersión; e) forma de distribución de la variable; y f) mecanismo de no respuesta. **Resultados:** los resultados muestran que todos los criterios de información funcionan bien en el escenario 1 para distribuciones normales y no normales con homogeneidad y heterogeneidad de varianzas. Sin embargo, en los escenarios 2 y 3, todos son precisos con la matriz ARH, aunque, AIC, AICCR y HQICR lo hacen para TOEP y UN. Por otro lado, cuando la distribución no es normal, solo en el escenario 3 funcionan bien AIC y AICCR, matrices más heterogéneas y No Estructurada, con Casos Completo MAR y MCAR. **Conclusiones:** en consecuencia, para seleccionar la matriz correctamente se recomienda analizar la heterogeneidad, tamaño muestral y distribución de los datos.

Palabras clave: criterios de información, estructuras de covarianza, pérdida datos, sensibilidad, diseños medidas repetidas.

Missing data affect data quality, analysis and subsequent decision making (Leke & Marwala, 2019). According to Dong and Peng (2013), missingness ranges in psychology and education studies fluctuate between 16% and 48% (Enders, 2003; Peng et al., 2006 in Dong & Peng, 2013) and, in repeated measures designs, Liu (2016) indicates that they occur in 96% of cases and are a major problem (Sullivan et al., 2017).

The inadequate treatment of missing data produces bias in the results (Liu, 2016) because there are smaller samples, more heterogeneous variances and unbalanced designs (Vallejo et al., 2019). This also affects the effect size, statistical power (Vallejo et al., 2019; Vallejo et al., 2018; Fernández et al., 2018; Zhang & Yuang, 2018), the predictive capacity of the model (Garson, 2020), the behavior of the information criteria (Vallejo et al., 2014; Livacic-Rojas et al., 2013; Livacic-Rojas et al., 2017) and the estimation of the parameters.

Focusing on the missingness process, Funatogawa and Funatogawa (2019) note that the mechanisms can be classified hierarchically on three levels: MCAR (the missingness mechanism is not dependent on the process), MAR (the missingness mechanism is dependent on observed responses and not on responses not observed) and NMAR (the missingness mechanism is dependent on responses not observed). Under MCAR, standard procedures provide consistent estimators and, when the mechanism is MAR, the missingness process does not need to be simultaneously

Received: February 19, 2020 • Accepted: May 27, 2020
Corresponding author: Pablo Livacic-Rojas
Faculty of Humanities - School of Psychology
University of Santiago de Chile
9170197 Santiago de Chile (Chile)
e-mail: pablo.livacic@usach.cl

modeled by probability methods because it can be factored in two parts (measurement and missingness processes). In this context, the maximum likelihood estimators (MLE) are consistent if the joint distribution of the response vector is correctly specified. In this regard, Little and Rubin (2020) point out that MLE are flexible, they avoid the use of ad hoc methods and they allow the estimation of variance even with missing data (see also in Molenberghs et al., 2015; Yoo, 2013). In the opposite scenario, generalized equation estimation (GEE) methods deliver biased estimators. Finally, under NMAR, MLEs provide biased estimators as they are the most complex mechanism, also affecting the prediction of patterns of change in study variables over time (Liu, 2016). In turn, under NMAR, the same author also notes that in longitudinal studies ignoring the missing data generates biased results, erroneous predictions and less mature statistical models than methods under MAR.

In an attempt to resolve the above, Funatogawa et al., (2019) indicate that in longitudinal designs procedures have been developed to analyze responses over time, the types of effect or covariates and their relationship. They also note that mutually dependent responses require methods that consider the structural configuration of the design based on correlations, the relationships between variances and covariances, the sample size, the number of measures and the relationship between them, because they influence the structures of means, variances, covariances, the (mixed or marginal) effects model and the method used for data analysis (maximum likelihood or other).

Vallejo et al., (2010) compared the efficacy of several information criteria to select nested covariance structures with the likelihood ratio test (LRT) under three different scenarios (maintaining constant matrix and mean adjustment; constant mean and matrix adjustment; mean and matrix adjustment), in repeated measures designs (q=6), with three covariance matrices [random coefficients (RC), first order auto-regressive heterogeneous (ARH1), unstructured (UN)], complete data, two groups (p=2; n_1=30; n_2=60), with positive and negative relationship between the matrices and group sizes, normal (γ_1= 0; γ_2=0) and non-normal distributions [Laplace (γ_1= 0; γ_2=3); exponential (γ_1= 2; γ_2=6) and lognormal (γ_1= 6.18; γ_2=110.94)] and assuming the data generating process to be true. They found that the selection criteria work better when the matrices are more complex, while LRT (based on full maximum likelihood, FML) works better than the rest, although it is less efficient when based on the restricted maximum likelihood method (REML).

In turn, Vallejo et al., (2011a) analyzed the effectiveness of several information criteria (AIC, AICc, BIC, CAIC, HQIC) in repeated measures designs (additive and non-additive models of four and eight measures), under three scenarios with mean models and covariance structures, (identical to the previous study), data with normal and non-normal distribution (exponential), four autoregressive covariance matrices [autoregressive (AR), ARH1, Toeplitz (TOEP) and UN], complete data and monotone missingness (MAR). Regardless of the estimation method used, the criteria work best when the group size and repeated measures increase. However, with missing data and lack of normality, their efficiency decreases, although it improves as the sample size increases.

With a slightly different scope to the two previous studies, Vallejo et al., (2011b) analyzed the error rates of type I and power by four methods [based on comparing the unstructured covariance structure, the true covariance structure (CPM-U vs. CPM-T), the multiple imputation method (based on generalized estimated equations) and the generalized equation method of weighted estimates (MI-GEE vs. WGEE)] for dealing with missing data in unbalanced repeated measures designs. They conclude that MI-GEE is the most robust for type I error rates with MAR and that CPM-T and CPM-U control these properly, while WGEE tends to inflate them. On the other hand, when missingness is NMAR, all show high type I error rates. With regard to statistical power, procedures based on covariance structures are clearly more powerful than MI-GEE and WGEE.

In the context of hierarchical and multigroup models, Vallejo et al. (2014) compared the effectiveness of AIC and BIC with other selectors (AICc, BIC, CAIC and HQIC) manipulating the intraclass correlation variables, number of groups, group size, parameter value and slope intercept covariance. They conclude that none of them functions correctly in all conditions or is consistently better than the others. Likewise, they indicate that AIC or AICC are more recommended when independent random effects are assumed, whereas BIC and AIC are more consistent when they are assumed dependent.

Next, in the context of the analysis of longitudinal data with incomplete measures and arbitrary covariance structures, Vallejo et al., (2018) report that with MAR, four repeated measures and homogeneity of variances, the CPM-U method (mixed linear model) controls type I error rates slightly better and exhibits higher power levels than the Brown-Forsythe procedures based on multiple imputations (MI-MBF), original data (OD-MBF) and complete cases (CC-MBF). However, if the homogeneity of variance is violated, MI-MBF better controls type I error rates (more conservative with four and six measures and liberal with eight) and shows similar power levels to the remaining procedures (except with OD-MBF and homogeneous variances). A similar pattern occurs with six and eight measures, although the control of type I error rates and power tends to worsen under conditions of heterogeneity.

Given the impact of the latter on different procedures, Vallejo et al., (2019) propose the solution of estimating the sample size necessary to reach power levels of 0.80 based on the ordinary least squares method and the empirical method of REML in multilevel designs for interaction effects with complete and incomplete data vectors (with homogeneous and heterogeneous variances at levels 1 and 2), relative bias levels (theoretical and empirical) and a MAR mechanism. The results show that if there is heterogeneity of variances and incomplete data vectors, larger sample sizes are required to reach the nominal power of 0.80, and to make more accurate estimates and more realistic conclusions.

Regarding the adjustment of various covariance structures using selection criteria under different conditions that can be generalized to multivariate contexts (Fernández et al., 2010), Livacic-Rojas et al., (2013) report that different studies show that, under different analytical conditions, AIC better selects the underlying structure of the data than BIC (48.13 versus 41.63% on average). However, a comparison of the frequency of selection of covariance structures (from 12 possible ones) and type I error rates between AIC and the correctly identified model (CIM) in a repeated measures design (p=3; q=4) with complete vectors, three covariance structures (RC, ARH, UN), three group sizes (30, 45, 60) and the relationship between the covariance structure and the group size (null, positive and negative) found that AIC selects the original covariance structure on 2% of occasions and that it does so on 48% of occasions for the heterogeneous version with slightly higher type I error rates.

Finally, associated with the previous study with similar manipulated variables, normal and non-normal distributions, complete vectors and truncated means, Livacic-Rojas et al., (2017) compared the sensitivity of AIC, MCI and MBF (Modified Brown Forsythe). They report that MBF functions better than AIC and CIM (correctly identified model) when the groups are small (n = 30), the relationship between covariance matrices and sample size is negative and there is a UN matrix for all types of effects.

Based on the background presented here, the objective of this paper is to analyze the effectiveness of different information criteria (AIC, AICc, BIC, CAIC and HQIC) to select the covariance structure when there are missing data in three different scenarios. It should be noted that with respect to previous studies, our study incorporates three missingness mechanisms (MAR, MCAR and NMAR), with a repeated measures design (p=2; q=5), different sample sizes (50, 100 and 120), homogeneity and heterogeneity of covariance matrices (1-3; 1-5) and their different types of relationship.

Method

A partially repeated measures design with two groups (p=2) and five measures (q=5) was used in a two-stage process. In the first stage, the following variables were manipulated: a) the model complexity; b) total sample size; c) equality, inequality and the type of dispersion matrices; d) the non-response mechanism and the type of missing data pattern. In the second stage, two more variables were added: e) heterogeneous matching of covariance matrices and sample size (null, positive and negative); f) non-normal distribution of data.

Variables manipulated in the two studies

Study one:

(a) Complexity of the model used to generate the data: a three-stage study was conducted. In the first stage, discriminating between MCAR, MAR and MNAR non-response patterns, the covariance structure was kept constant and the mean structure was adjusted. In the second stage, the mean structure was assumed known and the covariance structure was adjusted. In the third stage, both structures were adjusted at the same time. The following regression model was used in each of the studies: $\mathbf{E}\left(\boldsymbol{Y_{ijk}}\right)=\beta_0+\beta_1\boldsymbol{Trt_{ij}}+\beta_2\boldsymbol{Time_{ik}}+\beta_3\boldsymbol{Trt_{ij}}\times\boldsymbol{Time_{ik}}$, where y_{ijk} represents the response given by the *ith* subject $i = 1, \ldots, n_j$ of the *j-th* group on the *k-th* occasion $k = 1, \ldots, t_i$ Trt_{ij} denotes an indicator variable for the *ith* subject in the *j-th* treatment group and $Time_{ik}$ the corresponding time points. In addition, the adjustment of the mean structure required the generation of the data from an additive model (without $Trt_{ij} \cdot Time_{ik}$ interaction) and a non-additive model (with interaction). In both models, the magnitude of the regression coefficients was selected, attempting to ensure that the null hypotheses referring to fixed effects (complex case) or to the fixed design effect (simple case) were rejected in 80% of cases with a 95% confidence level. The value of the parameters that satisfy the aforementioned power was obtained using numerical techniques and the PROC MIXED module of the SAS program.

(b) Total sample size: For each *t* value, three total sample sizes $\boldsymbol{n}=\sum_j^g \boldsymbol{n_j}$ were considered (n = 50, n = 100 and n=120) as being representative of research in psychology or clinical trials.

(c) Pairing of covariance matrices and group size: When the design is balanced, the relationship between the size of the dispersion matrices and the group size is null. When the design is unbalanced, the relationship can be positive or negative. A positive relationship implies that the smallest group is associated with the smallest dispersion matrix, while a negative relationship implies that the smallest group is associated with the largest dispersion matrix.
For the different-sized groups the configuration was: (a) 20-30, 30-20, 10-40, 40-10 (n = 50); (b) 40-60, 60-40, 20-80, 80-20 (n = 100) and; (c) 50-70, 70-50, 30-90, 90-30 (n = 120).

(d) Equality of dispersion matrices: The performance of the selection criteria was evaluated with homogeneous and heterogeneous covariance matrices available in SAS and, in the first case, the elements of the two dispersion matrices were equal to each other $\Sigma_2=\Sigma_1$. In accordance with the work of Livacic-Rojas et al., (2013), the following matrices were used: random coefficients (RC) for Scenario 1, and first order autoregressive heterogeneous (ARH), Toeplitz (TOEP) and Unstructured (UN) for Scenarios 2 and 3.

(e) Non-response mechanism and type of missing data pattern:

1. MCAR with monotone missing data pattern**:** Two situations were specified. In the first, the MCAR process was denoted with a time-independent monotone missing data pattern (MCAR/IMP); subject *i* was not observed on either occasion *k* or the following occasions if $\boldsymbol{U_{ijk}} < \pi$. In the second, the MCAR process was denoted with a time-dependent monotone missing data pattern (MCAR/DMP); subject *i* was not observed on either occasion *k* or the following occasions if $\boldsymbol{U_{ijk}} < \pi_k$.
2. MAR with monotone missingness pattern: Two situations were specified. In the first, MAR was denoted with a monotone time-independent missingness pattern (MAR/IMP), subject *i* was not observed on occasion *k* or on the following occasions, if $\boldsymbol{U_{ijk}} < \varphi\ (\boldsymbol{Y_{ijk-1}} + \delta)$ where U_{ijk} is a uniformly distributed random variable. In this case, $\delta = \varphi^{-1}(\pi)2^{1/2}$, where $\varphi^{-1}(\cdot)$ is the inverse distribution or quantile function of the loss ratio π. In other words, $R_{ijk} = 0 \rightarrow R_{iju} = 0, \forall u > t$. In the second, MAR was denoted with a time-dependent monotone missingness pattern (MAR/DMP), subject *i* was not observed on occasion *t* or on the following occasions, if $U_{ijk} < \varphi(Yijk + \delta_k)$.
3. MNAR with monotone missing data pattern: Two situations were specified. In the first, MNAR was denoted with a monotone time-independent missingness pattern (MNAR/IMP); subject *i* was not observed on

occasion k or on subsequent occasions if $U_{ijk} < \varphi(Y_{ijk} + \delta)$. In the second, MNAR was denoted with a time-dependent monotone missingness pattern (MNAR/DMP); subject i was not observed on occasion k or on subsequent occasions if $U_{ijk} < \varphi(Y_{ijk} + \delta_k)$.

(f) Form of the distribution of the measurement variable: Although the MML is based on compliance with the normality assumption, when working with real data it is common for the indices of asymmetry (γ_1) and kurtosis (γ_2) to deviate from zero (Micceri, 1989), which may lead to an incorrect interpretation of the results. In this case, the population distributions was the multivariate normal distribution (γ_1 & $\gamma_2 = 0$). The following procedures were used in generating the data:

Normal case: In each treatment group, continuous longitudinal data were generated using the method of Ripley (1987). This procedure was performed in two steps:

1. Generation of z_{ij} pseudo-random observation vectors with $E(z_{ij}) = 0$ and $Cov(z_{ij})$= from a normal distribution, where **I is** the identity matrix. These vectors were obtained by means of the RANNOR function in SAS.
2. Creation of complete data sets y_{ij}, multiplying the vector z_{ij} by the Cholesky decomposition L_i, that is $y_{ij} = \beta_j + L_I z_{ij}$, where y_{ij} is a vector of length t for (i,jk)-th subject, $\mathbf{\beta}_j$ is a vector of dimension p containing the fixed effects of the population and $\boldsymbol{L_I}$ is a lower triangular matrix of dimension t satisfying $\Sigma_I = L_I L'I$, $I=1$,..6 or 12.

Study 2:

Equally, in addition to the variables a, b, c, d and e from study one, the following variables were manipulated:

(f) Inequality of dispersion matrices: The elements of one of the matrices were three and five times greater than those of the other $\Sigma_2 = 5\Sigma_1$ with the same conditions and matrices as indicated in point d.

(g) Form of the distribution of the measurement variable: In this case, the population distributions was a moderately skewed distribution with parameters equivalent to nonnormal distributions ($\gamma1 = 4$; $\gamma2 = 42$), Laplace ($\gamma1 = 2$; $\gamma2 = 6$), log normal ($\gamma1 = 1.7501897$; $\gamma2 = 5.8984457$) and exponential ($\gamma1 = 6.1848771$; $\gamma2 = 110.93639$). In this analytical context, the non-normal distribution with ($\gamma1 = 4$; $\gamma2 = 42$) has been used to assess the impact that moderate bias and peak indexes would have on the performance of the five information criteria.

The following procedures were used in generating the data:

Non-normal case: Non-normal multivariate data were generated using the power method developed by Fleishman (1978) and extended to multivariate situations by Vale and Maurelli (1983). This procedure involved the following steps:

1. Calculating a weight vector w = $[a\ b\ c\ d]'$ with the desired indices of asymmetry and kurtosis for each distribution, using the Fleishman power method.
2. Calculating an appropriate intermediate correlation matrix, R_I, solving for all possible pairs of repeated measures with the following third order polynomial equation: $R_{xkxk'} = \rho_{ZkZk'}(b^2 + 6bd + 9d^2) + \rho^2_{ZkZk'}2c^2 + \rho^3_{ZkZk'}6d^2$, where $\rho_{ZkZk'}$ is the correlation coefficient between two standard normal variables and $X_k = (a + bZ_k + cZ^2_k + dZ^3_k)$ and $X_{k'} = (a + bZ_{k'} + cZ^2_{k'} + dZ^3_{k'})$ are the two correlated non-normal variables.
3. Factoring the intermediate correlation matrix R_I to generate a vector of multivariate random normal variables with the prescribed $\rho_{ZkZk'}$, that is, $x_{ij} = M_I z_{ij}$ where $\mathbf{x}_{ij}$ denotes the vector of variables transformed with $E(x_{ij}) = 0$ and $Cov(x_{ij}) = R_I$ and $\mathbf{M}_I$ is the lower triangular matrix obtained by Cholesky decomposition, with the property $R_I = M_I M'_I$.
4. Transforming the variables generated in the previous step so that they take the desired distribution form, as well as the desired fixed effects and variances, that is, $y_{ij} = \beta_j + D_I(X'_{ij}w)$, where $\mathbf{D}_I$ is a diagonal matrix containing the standard deviations of the covariance matrix Σ_I and $X'_{ij} = [1_k x_{ij} x^2_{ij} x^3_{ij}]$. In the line of studies about non normal distributions, the reader can also consult Blanca et al., (2013) and Bono et al., (2017; 2020).

Finally, the information criteria used are described in the works of Livacic-Rojas et. al., (2013; 2017) and Vallejo et al., (2014).

Results

In study 1, the results of Table 1 (normal distribution with homogenous relationship between the size of the group and 1-1 variances) in scenario 1 (RC matrix) show high power levels (80% or higher, Aberson, 2019) for the five information criteria in all the analyzed conditions. However, in scenario 2 under ARH matrix, higher power is observed for HQICR, BICR and CAICR criteria (60% of the analyzed conditions) for all the missing data mechanisms. In turn, under TOEP matrix, the AIC and AICCR criteria are efficient in 35% of the analyzed conditions, whereas under UN matrix these work well in 15% of the occasions for the Complete Cases and MCAR missingness mechanisms. In scenario 3, under ARH matrix, power is high in 60% of the conditions analyzed for HQICR, BICR and CAICR criteria with all the missingness mechanisms. In turn, under TOEP and UN, power levels are high in 55% and 35% of the conditions with AIC and AICCR criteria in all mechanisms.

In turn, Table 2 (normal distribution, with group heterogeneous relationships and 1-3; 1-5 variances) shows values similar to the table above, whereas power is slightly higher when the heterogenous relationship is 1-5. In scenario 2, all the criteria show high power in ARH matrix in 93% of the conditions analyzed with all the mechanisms. In turn, under TOEP matrix, AIC and AICCR criteria work well in 23% of the conditions analyzed with Complete Cases, MAR and MCAR mechanisms and the two heterogeneous relationships. On the other hand, in scenario 3, under ARH, the HQICR, BICR and CAICR criteria are efficient with all the missingness mechanisms in 50% of the conditions and the two heterogeneous relationships. However, in TOEP and UN matrices, the AIC and AICCR criteria work better, with all the missingness mechanisms and the two heterogeneous relationships in 20% and 38% of the conditions, respectively.

On the other hand, in study 2, when the distribution is non-normal (γ_1=4; γ_2=42 and, with heterogeneous relationships between groups and 1-3; 1-5 variance), Table 3 shows high power only in scenario 1 (RC matrix) in 98% of the analyzed conditions (under RC matrix) and associated to all the information criteria, missingness mechanisms and heterogeneous relationships. In scenario 2, criteria are inefficient in 100% of reported conditions. Finally, in scenario 3, under the UN matrix, the AIC and AICCR criteria are efficient in 20% of the occasions with Complete, MCAR, NMAR criteria and 1-5 heterogenous relationship.

In turn, table 4 (Laplace distribution), shows in scenario 1 (RC matrix) high power in 97.5% of the conditions for all the information criteria, missingness mechanisms and heterogeneous relationships. In scenario 2, criteria are inefficient in 100% of the reported conditions. In turn, in scenario 3, under ARH matrix, power is high in 2.5% of the conditions with CAICR criterion, the MCAR mechanism and 1-5 heterogeneous relationship. Subsequently, under UN matrix, power is high in 27.5% of the conditions with AIC and AICCR criteria, the Complete Cases mechanisms, MCAR, NMAR and the two heterogeneous relationships.

In turn, table 5 (normal Log distribution), in scenario 1 (RC matrix), the criteria are efficient in 100% of reported conditions. In scenario 2, under ARH matrix, power is high in 30% of the conditions analyzed with HQICR, BICR and CAICR criteria, all the missingness mechanisms and the two heterogeneous relationships. Under TOEP, power is high only in 2.5% of the conditions with AIC criterion, the Complete Cases mechanism and 1-5 heterogeneous relationship. In scenario 3, associated to ARH, power is high only in 2.5% of the conditions with HQICR, BICR, CAICR criteria, all the Complete Cases missingness mechanisms and the two heterogeneous relationships. For the UN matrix, power is high in 30% of the conditions with AIC, AICCR criteria, all Complete Cases missingness mechanisms and the two heterogeneous relationships.

Finally, Table 6 (Exponential distribution) shows high power in scenario 1 (RC matrix) in all conditions analyzed for all criteria, mechanisms and the two heterogeneous relationships. In scenario 2, the criteria are inefficient in 100% of reported conditions. In scenario 3, power is high in 17.5% of conditions analyzed under UN matrix with AIC and AICC criteria, Complete Cases, MCAR and the two heterogenous relationships.

Discussion

In the present study the sensitivity levels of five information criteria have been analyzed to select covariance structures when there are different missingness mechanisms in three different scenarios.

In study 1, scenario 1 (maintaining constant matrix and mean adjustment) the five information criteria work well to discriminate the covariance structure for missing data with the different mechanisms. In scenarios 2 (constant mean and matrix adjustment) and 3 (constant mean and matrix), a similar situation is observed, but with higher efficiency under ARH matrix. Subsequently, when the relationships between groups and variances are heterogeneous, in

Table 1

Average statistical power levels for samples of 50, 100 and 120 cases in three different scenarios with normal distribution (γ_1=0; γ_2 =0) and relationship of groups with homogeneity of variance (1-1)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-1	**95.1**	71.5	**80.7**	**89.8**	76.5	**89.8**	**90.3**
	AICCR	1-1	**95.1**	75.9	**84.1**	**87.3**	78.8	**90.7**	**87.7**
	HQICR	1-1	**94.5**	**95.8**	**81.2**	65.2	**95.7**	**82.7**	65.3
	BICR	1-1	**94.5**	**98.6**	**80.4**	44.6	**99.0**	**80.7**	44.7
	CAICR	1-1	**93.5**	**98.8**	68.2	25.6	**99.7**	71.6	25.7
MAR	AIC	1-1	**95.4**	69.1	76.7	79.4	76.0	**82.8**	**80.4**
	AICCR	1-1	**95.4**	73.2	78.4	75.5	78.5	**83.1**	76.7
	HQICR	1-1	**95.3**	**91.2**	76.5	56.2	**92.7**	77.9	57.3
	BICR	1-1	**95.4**	**98.5**	57.2	21.1	**98.8**	56.1	21.8
	CAICR	1-1	**94.7**	**99.4**	41.9	8.34	**99.5**	40.9	9.19
MCAR	AIC	1-1	**95.3**	69.2	**80.1**	**82.0**	76.1	**85.5**	**85.5**
	AICCR	1-1	**95.3**	73.1	**82.1**	78.9	78.5	**86.1**	**82.1**
	HQICR	1-1	**95.2**	**91.7**	**83.0**	63.3	**92.8**	**83.8**	65.7
	BICR	1-1	**94.7**	**98.8**	64.9	31.8	**98.9**	65.4	30.8
	CAICR	1-1	**93.6**	**99.1**	53.6	17.8	**99.3**	53.5	18.7
NMAR	AIC	1-1	**95.6**	68.5	76.8	79.5	76.1	**83.7**	**80.8**
	AICCR	1-1	**95.6**	72.1	78.5	76.8	78.2	**83.8**	77.8
	HQICR	1-1	**95.5**	**91.6**	77.4	57.0	**92.9**	78.8	56.1
	BICR	1-1	**94.8**	**98.5**	55.5	21.4	**98.7**	58.7	21.6
	CAICR	1-1	**93.8**	**99.5**	41.5	8.83	**99.6**	44.8	8.93

Legend: mechanism (M); complete case (C); missing at random (MAR); missing completely at random (MCAR); not missing at random (NMAR); IC (information criterion); AIC (Akaike information criterion); AICCR (Akaike information criterion, corrected robust); BICR (Bayesian information criterion, robust); HQICR (Hannan-Quinn information criterion, robust); CAICR (consistent Akaike information criterion, robust); HBGM (homogeneous relation between group matrices 1-1); Scenario 1 (S1); Scenario 2 (S2); Scenario 3 (S3); random coefficients (RC); autoregressive heterogeneity (ARH); Toepliz (TOEP); unstructured (UN); γ_1= Skewness; γ_2 =Kurtosis

Table 2
Average statistical power levels for samples of 50, 100 and 120 with normal distribution (γ_1=0; γ_2 =0) and relationship groups-heterogeneity of variance (1-3; 1-5)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-3	**90.1**	**84.2**	**88.7**	75.0	64.2	**81.8**	**90.8**
		1-5	**85.8**	**84.2**	**89.6**	76.6	53.3	**83.8**	**88.4**
	AICCR	1-3	**90.1**	**87.8**	**84.2**	63.4	67.1	**83.8**	**88.4**
		1-5	**85.8**	**87.9**	**86.2**	66.7	56.0	75.8	**88.4**
	HQICR	1-3	**90.0**	**95.4**	72.5	35.6	**86.2**	**88.6**	77.6
		1-5	**85.5**	**97.2**	74.1	38.6	78.5	84.7	79.3
	BICR	1-3	**89.3**	**92.9**	39.2	2.85	**96.8**	78.3	48.7
		1-5	**85.3**	**99.6**	39.7	3.85	**93.6**	77.1	50.9
	CAICR	1-3	**88.2**	**86.3**	17.6	0.21	**99.0**	67.3	30.1
		1-5	**83.4**	**99.2**	19.3	0.34	**97.3**	66.7	33.7
MAR	AIC	1-3	**91.1**	**85.6**	**88.8**	67.5	61.4	79.3	**89.1**
		1-5	**84.5**	**84.6**	**87.0**	65.2	46.8	65.6	**90.3**
	AICCR	1-3	**91.1**	**88.8**	**86.1**	56.3	63.3	**81.1**	**87.0**
		1-5	**84.5**	**88.4**	**84.4**	52.4	48.9	68.5	**88.5**
	HQICR	1-3	**91.1**	**98.6**	56.5	18.2	**85.7**	**82.9**	66.2
		1-5	**84.4**	**98.8**	52.5	18.8	76.3	77.8	71.1
	BICR	1-3	**91.1**	**99.2**	6.30	0.34	**97.9**	61.7	30.3
		1-5	**84.3**	**100**	8.10	0.35	**94.1**	58.0	36.3
	CAICR	1-3	**91.1**	**96.8**	0.90	0.00	**99.0**	41.3	13.5
		1-5	**84.1**	**100**	1.80	0.00	**97.3**	39.5	19.1
MCAR	AIC	1-3	**89.0**	79.6	79.8	68.7	63.8	77.4	**86.9**
		1-5	**86.0**	**82.2**	**81.6**	70.4	52.2	69.2	**88.2**
	AICCR	1-3	**88.9**	**82.6**	74.7	55.6	66.5	79.1	**83.8**
		1-5	**86.0**	**87.1**	76.5	57.9	55.5	71.5	**85.7**
	HQICR	1-3	**88.8**	**88.7**	57.8	25.0	**85.5**	**80.5**	68.8
		1-5	**85.8**	**96.4**	59.5	28.1	77.4	76.9	71.4
	BICR	1-3	**88.1**	**96.4**	36.8	16.9	**92.1**	**83.6**	50.9
		1-5	**85.1**	**98.6**	18.9	1.59	**93.1**	63.2	40.0
	CAICR	1-3	**87.1**	75.7	4.19	0.05	**98.9**	50.7	19.0
		1-5	**84.2**	**97.3**	4.85	0.08	**97.1**	50.2	22.8
NMAR	AIC	1-3	**88.1**	**80.3**	75.7	62.6	63.2	74.6	**82.6**
		1-5	**85.9**	**81.4**	77.0	64.0	50.7	65.1	**84.5**
	AICCR	1-3	**88.0**	**84.8**	70.7	49.2	66.4	75.9	78.9
		1-5	**86.0**	**85.8**	71.6	50.9	54.0	67.2	**81.3**
	HQICR	1-3	**88.0**	**92.3**	48.7	18.7	85.5	75.4	60.9
		1-5	**85.7**	**96.1**	51.1	21.0	76.6	71.6	64.7
	BICR	1-3	**87.3**	**86.8**	9.68	0.47	**96.7**	55.3	26.5
		1-5	**84.9**	**98.4**	10.8	1.01	**92.4**	55.0	32.1
	CAICR	1-3	**86.3**	77.4	1.85	0.00	**98.7**	40.4	12.4
		1-5	**83.7**	**97.1**	2.40	0.08	**96.8**	40.4	16.7

Legend: mechanism (M); IC (information criterion); HBGM (heterogeneous relation between group matrices 1-3; 1-5); complete case (C); missing at random (MAR); missing completely at random (MCAR); not missing at random (NMAR); AIC (Akaike information criterion); AICCR (Akaike information criterion, corrected robust); BICR (Bayesian information criterion, robust); HQICR (Hannan-Quinn information criterion, robust); CAICR (corrected Akaike information criterion, robust); Scenario 1 (S1); Scenario 2 (S2); Scenario 3 (S3); random coefficients (RC); autoregressive heterogeneity (ARH); Toepliz (TOEP); unstructured (UN); γ_1= Skewness; γ_2 =Kurtosis

scenario 2, AIC shows slightly higher power than the other criteria, whereas in scenario 3 these are observed in association with HQICR, BICR and CAICR, and all the missingness mechanisms. These results coincide with those of Vallejo et., (2011a; 2014; 2019) and Livacic-Rojas et al., (2013; 2017). Notwithstanding the foregoing, it is important for researchers to take these results with caution because heterogeneous variances impact negatively, increasing the error, which affects parameter estimation (Raghunathan, 2016). Moreover, the results of Scenario 1 are unrealistic in psychology research that uses repeated measures designs in applied settings.

Table 3

Average statistical power levels for samples of 50, 100 and 120 with non-normal (γ_1=4; γ_2 =42) and relationship groups-heterogeneity of variance (1-3; 1-5)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-3	**97.0**	19.8	57.8	74.9	17.8	29.2	**86.9**
		1-5	**86.9**	15.8	59.8	77.6	14.7	24.0	**88.5**
	AICCR	1-3	**97.0**	23.3	55.7	64.4	19.2	30.5	**83.3**
		1-5	**86.8**	18.9	58.5	68.5	16.0	26.1	**86.1**
	HQICR	1-3	**96.7**	37.4	49.2	41.6	32.3	42.9	71.6
		1-5	**86.7**	34.9	53.6	46.3	27.8	39.1	74.8
	BICR	1-3	**96.2**	55.2	29.8	12.7	46.9	52.6	51.1
		1-5	**85.3**	61.0	30.1	11.5	45.8	48.6	48.6
	CAICR	1-3	**94.1**	62.6	12.5	2.35	60.0	47.8	31.3
		1-5	**83.9**	71.4	16.2	4.15	55.0	46.9	34.9
MAR	AIC	1-3	**86.4**	22.6	48.8	66.3	22.0	30.0	78.7
		1-5	**86.8**	24.0	50.2	68.7	19.2	26.0	79.7
	AICCR	1-3	**86.9**	26.5	46.1	54.0	23.8	31.4	75.4
		1-5	**86.7**	28.4	47.4	57.2	20.8	27.9	76.6
	HQICR	1-3	**86.0**	42.0	34.1	29.9	38.8	41.6	58.0
		1-5	**86.5**	46.8	35.0	31.0	35.3	37.3	60.5
	BICR	1-3	**84.9**	56.0	8.63	5.33	58.3	41.1	31.2
		1-5	**85.4**	69.0	8.56	4.80	54.7	36.9	32.4
	CAICR	1-3	**83.7**	56.7	1.99	2.07	67.7	34.0	16.4
		1-5	**84.0**	74.0	2.18	0.85	64.8	29.4	19.3
MCAR	AIC	1-3	**84.3**	16.8	48.3	70.7	18.4	28.4	**83.3**
		1-5	**86.1**	18.0	51.3	71.7	15.5	23.7	**83.8**
	AICCR	1-3	**79.7**	19.3	42.5	55.4	18.8	28.8	75.3
		1-5	**86.0**	21.5	49.1	60.6	16.8	25.5	**80.8**
	HQICR	1-3	**84.1**	35.3	36.1	37.8	33.9	41.5	65.6
		1-5	**85.7**	38.0	40.0	36.5	29.3	36.0	64.9
	BICR	1-3	**83.2**	54.8	11.8	10.8	52.7	43.5	37.5
		1-5	**84.6**	62.2	14.3	7.07	47.5	39.4	37.5
	CAICR	1-3	**81.2**	59.6	3.58	5.31	62.5	37.2	23.6
		1-5	**83.3**	71.9	4.74	1.82	57.4	34.5	23.7
NMAR	AIC	1-3	**84.2**	21.0	46.5	65.4	21.8	31.0	79.9
		1-5	**86.8**	23.4	50.0	68.3	18.6	25.8	**81.1**
	AICCR	1-3	**84.1**	24.9	44.6	54.2	23.4	32.7	76.5
		1-5	**86.8**	27.7	47.6	57.1	20.0	27.6	78.1
	HQICR	1-3	**83.8**	39.9	32.5	28.1	38.1	42.5	59.7
		1-5	**86.6**	43.4	35.5	32.1	33.8	37.5	61.9
	BICR	1-3	**82.8**	56.6	7.56	3.68	57.1	40.5	30.5
		1-5	**85.6**	63.2	9.05	4.92	52.5	37.6	34.4
	CAICR	1-3	**81.6**	58.6	1.79	0.59	66.7	32.3	17.6
		1-5	**84.2**	70.7	2.20	0.91	61.6	30.3	20.6

Legend: Same as Table 2

On the other hand, there are two different analytical contexts in study 2. On one hand, with (non-normal with γ_1=4; γ_2=42, Lapalce and Exponential) distributions, power is high in scenario 1 for most information criteria, missingness mechanisms and two heterogenous relationships. In turn, in scenario 3 under the UN matrix, the AIC and AICCR criteria are efficient with Complete Cases, MAR, MCAR and NMAR mechanisms. On the other hand, with Log normal distribution, in scenario 1 the situation is similar to previous distributions. In scenario 2, under ARH matrix, power is high with BICR and CAICR criteria, all missingness criteria and heterogenous relationships. In scenario 3, associated to UN, there is high power with AIC and AICCR, all missingness mechanisms and heterogeneous relationships. These results coincide in part with those of Vallejo et. al., (2014; 2018; 2019); Livacic Rojas et. al., (2013; 2017), but are in the opposite direction to the indications of Vallejo et. al., (2011a; 2011b) since the greater the heterogeneity,

Table 4
Average statistical power levels for samples of 50, 100 and 120 with Laplace (γ_1 =2; γ_2 = 6) and relationship groups-heterogeneity of variance (1-3; 1-5)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-3	**86.0**	19.3	54.4	72.7	17.0	9.92	**86.6**
		1-5	**88.2**	14.8	58.4	69.4	14.1	22.3	**87.9**
	AICCR	1-3	**81.3**	20.4	49.6	60.5	16.6	9.86	80.0
		1-5	**88.2**	17.9	55.5	61.3	15.5	24.4	**85.8**
	HQICR	1-3	**85.7**	33.9	46.5	42.4	29.9	12.6	73.2
		1-5	**87.9**	32.0	49.5	43.8	26.0	36.5	72.8
	BICR	1-3	**84.8**	53.0	25.3	9.71	47.5	13.7	47.4
		1-5	**86.8**	56.1	27.2	17.2	41.1	46.3	46.8
	CAICR	1-3	**83.3**	59.0	12.0	2.78	57.6	12.4	32.7
		1-5	**85.2**	67.7	13.1	11.6	49.2	44.7	32.4
MAR	AIC	1-3	**85.5**	18.9	48.0	67.5	20.7	25.4	**80.2**
		1-5	**87.7**	21.2	48.3	68.1	17.7	24.5	**80.8**
	AICCR	1-3	**85.4**	22.4	45.5	55.3	22.3	27.4	77.0
		1-5	**88.6**	25.9	46.1	57.3	18.8	26.4	78.1
	HQICR	1-3	**85.1**	35.9	35.3	29.9	36.9	37.0	60.4
		1-5	**88.3**	43.6	35.6	32.8	32.5	36.4	61.7
	BICR	1-3	**83.9**	52.1	11.8	4.13	55.9	37.9	31.9
		1-5	**87.0**	65.3	10.1	5.15	51.4	37.1	34.2
	CAICR	1-3	**82.2**	54.6	3.92	0.72	65.2	31.1	18.2
		1-5	**85.4**	72.3	2.88	1.19	61.0	31.1	20.6
MCAR	AIC	1-3	**85.7**	15.0	45.8	69.8	16.7	25.8	**82.7**
		1-5	**88.8**	15.5	48.1	72.9	38.7	21.8	84.3
	AICCR	1-3	**85.7**	18.2	43.7	58.1	18.1	28.1	79.6
		1-5	**88.7**	19.2	46.0	63.8	41.5	23.6	81.6
	HQICR	1-3	**85.5**	31.8	34.4	33.2	30.0	39.2	65.0
		1-5	**88.4**	34.4	37.7	39.5	60.7	33.6	68.1
	BICR	1-3	**84.4**	51.0	11.4	5.59	49.0	41.1	37.7
		1-5	**87.3**	58.9	13.9	8.66	77.8	37.8	40.8
	CAICR	1-3	**83.0**	56.7	3.41	1.22	58.5	35.3	23.1
		1-5	**85.7**	69.5	4.42	2.78	83.2	33.4	26.9
NMAR	AIC	1-3	**86.3**	18.3	46.6	68.4	19.9	28.2	**80.5**
		1-5	**86.2**	20.9	47.8	69.0	17.5	25.8	**81.6**
	AICCR	1-3	**86.2**	22.2	43.6	57.0	21.6	30.1	77.1
		1-5	**85.7**	22.7	46.2	57.4	18.8	26.5	78.8
	HQICR	1-3	**85.0**	36.9	31.2	32.2	35.6	39.9	60.9
		1-5	**85.5**	37.6	35.6	32.2	31.8	35.5	63.8
	BICR	1-3	**82.0**	54.1	7.25	6.64	54.0	40.2	32.3
		1-5	**84.4**	58.7	9.59	5.53	50.0	37.3	35.6
	CAICR	1-3	79.5	58.2	1.65	2.49	63.5	32.8	18.9
		1-5	**83.2**	66.8	2.58	1.20	60.1	30.7	21.8

Legend: Same as Table 2

the lower the power levels expected. They should also be taken with caution because the missingness pattern considered is monotone, which might not be the case if this situation were to change.

For applied studies, researchers are advised that if they detect missing data, they should use the information criteria to detect the covariance structure that underlies the data in order to apply more appropriate methods for parameter estimation. Regarding this point, Leke et al., (2019) note that imputation methods (individual or multiple) are very viable in the social sciences. However, it is important to consider that multivariate methods tend to assume that the missingness mechanisms are MCAR or MAR with monotone patterns (Little & Rubin, 2020) and that it is uncommon for the missingness mechanism to be estimated with some degree of confidence (Funatogawa et al., 2019). If the mechanism were NMAR, biased results and wrong predictions would be expected to be produced (Liu, 2016). On a similar note, since the structure of

Table 5
Average statistical power levels for samples of 50, 100 and 120 with log normal distribution (γ_1 =1.7501897; γ_2 = 5.8984457) and relationship groups-heterogeneity of variance (1-3; 1-5)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-3	**84.7**	47.5	77.2	75.4	35.1	51.9	**89.6**
		1-5	**85.6**	43.8	**81.9**	76.5	28.3	43.8	**89.8**
	AICCR	1-3	**84.7**	50.8	74.2	64.8	37.4	55.2	**87.2**
		1-5	**85.6**	50.0	78.7	65.6	30.9	46.9	**87.4**
	HQICR	1-3	**84.6**	71.3	65.9	40.9	57.2	70.1	75.5
		1-5	**85.3**	72.7	69.1	41.0	49.5	64.3	75.3
	BICR	1-3	**83.9**	**82.9**	38.3	11.2	73.1	67.5	47.9
		1-5	**84.5**	**91.6**	38.1	6.64	72.2	70.3	47.8
	CAICR	1-3	**83.0**	**82.5**	19.5	2.97	**86.0**	63.7	30.4
		1-5	**83.4**	**95.3**	19.9	1.12	**81.4**	64.3	31.1
MAR	AIC	1-3	**86.0**	50.6	69.3	63.8	40.7	53.2	**80.2**
		1-5	**84.5**	53.0	69.6	67.5	33.7	44.7	**81.5**
	AICCR	1-3	**86.0**	55.8	65.0	50.7	43.2	55.7	76.5
		1-5	**84.4**	59.4	65.5	55.7	36.0	47.3	77.7
	HQICR	1-3	**85.8**	73.6	47.0	22.0	64.5	64.2	58.1
		1-5	**84.2**	**80.0**	47.7	30.8	56.9	57.1	59.5
	BICR	1-3	**85.2**	77.3	11.0	1.08	**83.5**	53.9	26.9
		1-5	**83.4**	**91.9**	11.1	6.86	78.3	49.8	27.9
	CAICR	1-3	**84.1**	70.0	2.46	0.08	**90.1**	41.6	13.6
		1-5	**82.2**	**91.5**	2.60	3.16	**86.3**	38.4	14.7
MCAR	AIC	1-3	**85.3**	44.1	70.5	69.1	35.3	49.5	**85.0**
		1-5	**84.3**	44.6	72.4	70.6	30.5	42.7	**86.4**
	AICCR	1-3	**85.3**	49.9	66.5	56.8	37.9	50.6	**81.8**
		1-5	**84.3**	51.0	68.6	58.2	33.0	44.6	**83.7**
	HQICR	1-3	**85.1**	70.4	51.4	28.5	58.6	62.0	67.1
		1-5	**84.0**	73.6	53.5	31.2	53.4	56.5	69.1
	BICR	1-3	**84.3**	**80.6**	15.2	2.30	79.3	56.5	36.2
		1-5	**83.2**	**91.5**	17.2	3.12	74.9	53.5	39.8
	CAICR	1-3	**83.0**	76.8	4.14	0.15	**87.2**	46.1	20.9
		1-5	**82.0**	**93.5**	4.66	0.38	**83.2**	44.1	23.9
NMAR	AIC	1-3	**85.3**	48.7	64.4	64.5	39.6	52.0	**81.6**
		1-5	**84.9**	52.2	69.4	67.1	33.2	43.3	**82.8**
	AICCR	1-3	**85.3**	54.4	60.4	52.1	42.3	54.4	78.1
		1-5	**84.9**	58.9	65.2	54.5	36.0	45.1	79.7
	HQICR	1-3	**85.1**	73.1	45.2	22.8	63.1	63.1	60.4
		1-5	**84.8**	79.1	47.1	26.6	56.6	55.6	63.6
	BICR	1-3	**84.4**	79.9	14.2	1.39	**82.3**	52.2	28.9
		1-5	**84.0**	**92.5**	11.0	2.17	77.8	48.6	33.3
	CAICR	1-3	**83.3**	74.2	7.88	0.08	**89.3**	39.0	14.8
		1-5	**82.9**	**93.6**	2.78	0.25	**85.3**	37.8	18.3

Legend: Same as Table 2

single-level data hinders appropriate analysis with missing values (Heck & Thomas, 2020), it is recommended to use multilevel models with larger sample sizes in order to have more efficient estimators (Vallejo et al., 2019).

Finally, it would be appropriate for future studies to analyze the behavior of different information criteria in selecting covariance structures, considering the percentage of missing data, the effect size, larger samples, the relationship between the matrices and the more heterogeneous groups (Grissom et al., 2012), and other non-normal distributions.

Acknowledgements

The authors would like express acknowledge the suggestions of revisors to improve this work has been funded by the Chilean National Fund for Scientific and Technological Development (FONDECYT. Ref.: 1170642) and the Spanish Ministry of Science, Innovation and Universities (Ref: PGC2018-101574-B-I00).

Table 6
Average statistical power levels for samples of 50, 100 and 120 with exponential (γ_1 =6.1848771; γ_2 = 110.93639) and relationship groups-heterogeneity of variance (1-3; 1-5)

			S1	S2			S3		
M	**IC**	**HBGM**	**RC**	**ARH**	**TOEP**	**UN**	**ARH**	**TOEP**	**UN**
C	AIC	1-3	**88.4**	5.36	31.7	76.1	8.33	14.3	**85.5**
		1-5	**92.2**	5.07	34.2	76.5	7.13	12.8	**83.1**
	AICCR	1-3	**88.4**	6.65	30.5	67.6	9.21	15.9	**83.1**
		1-5	**92.1**	6.52	33.1	68.4	7.86	13.8	**81.8**
	HQICR	1-3	**87.4**	13.1	26.4	49.1	16.1	24.3	70.8
		1-5	**91.8**	13.3	30.1	50.9	14.3	21.3	70.7
	BICR	1-3	**87.0**	25.3	15.3	17.3	27.8	33.1	48.1
		1-5	**90.6**	27.6	18.8	20.4	25.0	28.0	48.4
	CAICR	1-3	**85.5**	32.8	7.65	7.90	35.2	33.1	35.0
		1-5	**89.1**	38.0	9.94	10.0	32.1	29.5	35.8
MAR	AIC	1-3	**85.7**	9.27	30.0	69.3	11.6	15.6	78.0
		1-5	**93.7**	10.5	30.3	70.1	10.6	13.7	78.6
	AICCR	1-3	**85.6**	11.2	28.6	59.0	12.7	16.7	75.0
		1-5	**93.6**	12.7	27.9	60.5	11.4	15.1	75.9
	HQICR	1-3	**85.3**	19.4	21.4	36.1	22.1	23.7	59.0
		1-5	**93.3**	22.5	20.1	38.0	20.3	21.4	60.1
	BICR	1-3	**84.2**	31.2	6.41	9.68	35.9	26.4	35.3
		1-5	**92.1**	39.6	5.33	9.91	34.1	22.7	35.7
	CAICR	1-3	**82.5**	36.4	1.70	2.59	45.3	22.2	21.8
		1-5	**90.4**	48.1	1.92	3.46	43.0	18.7	23.9
MCAR	AIC	1-3	**89.9**	5.37	26.5	72.6	9.12	14.3	**82.9**
		1-5	**94.1**	5.90	30.8	71.7	7.74	12.0	**81.8**
	AICCR	1-3	**89.9**	6.94	25.6	63.4	9.95	15.5	**80.6**
		1-5	**94.0**	7.65	29.0	62.4	8.64	13.2	79.2
	HQICR	1-3	**89.5**	14.2	20.6	40.9	17.6	22.6	66.3
		1-5	**93.5**	15.4	23.2	42.4	15.2	20.0	66.2
	BICR	1-3	**88.3**	27.7	7.89	10.9	30.1	27.0	41.4
		1-5	**92.3**	31.0	9.43	12.7	27.0	24.8	42.7
	CAICR	1-3	**87.1**	35.6	2.96	4.06	37.9	23.8	28.8
		1-5	**90.4**	41.4	3.62	4.99	35.3	22.7	30.1
NMAR	AIC	1-3	**89.6**	7.75	24.5	69.6	11.4	16.3	78.0
		1-5	**93.8**	8.09	28.2	69.3	9.87	19.2	74.0
	AICCR	1-3	**89.6**	9.58	23.3	58.7	12.3	17.7	75.1
		1-5	**93.8**	9.99	27.0	61.0	10.6	20.2	71.8
	HQICR	1-3	**89.3**	17.6	18.1	37.3	20.5	24.7	59.2
		1-5	**93.4**	19.0	20.7	38.0	18.4	25.1	57.8
	BICR	1-3	**88.2**	31.9	6.57	10.6	33.9	26.4	33.8
		1-5	**92.1**	36.6	5.86	9.99	31.2	25.2	36.2
	CAICR	1-3	**86.3**	37.2	3.62	4.58	42.6	22.0	21.9
		1-5	**89.5**	37.8	7.52	15.3	34.0	19.7	35.3

Legend: Same as Table 2

References

Aberson, C.L. (2019). *Applied Power Analysis for the Behavioral Sciences*. Routledge.

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 9*(2), 78-84. https://doi.org/10.1027/1614-2241/a000057.

Bono, R., Blanca, M. J., Arnau, J., & Gómez-Benito, J. (2017). Non-normal distributions commonly used in health, education, and social sciences:

A systematic review. *Frontiers in Psychology-Quantitative Psychology and Measurement, 8*, Article 1602, 1-6. https://doi.org/10.3389/fpsyg.2017.01602

Bono, R., Arnau, J., Alarcón, R., & Blanca, M. J. (2020). Bias, precision and accuracy of skewness and kurtosis estimators for frequently used continuous distributions. *Symmetry, 12*(1), 19. https://doi.org/10.3390/sym12010019

Dong & Peng (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*, 222. https://doi.org/10.1186/2193-1801-2-222

Fernández, P., Vallejo G., & Livacic-Rojas P. (2010). Robustez de cinco estadísticos univariados para analizar diseños Split Plot en condiciones adversas [Robustness of five univariate statistics to analyze Split Plot designs in adverse conditions]. *Revista Latinoamericana de Psicología, 42(2),* 289-309.

Fernández, P., Vallejo, G., Livacic-Rojas, P., & Tuero-Herrero, E. (2018). The (Ir)Responsibility of (Under)Estimating Missing Data. *Frontiers in Psychology. 9,* 556-568. https://doi.org/10.3389/fpsyg.2018.00556

Fleishman, A.I. (1978). A method for simulating non-normal distributions. *Psychometrika*, *43*, 521-532. https://doi.org/10.1007/BF02293811

Funatogawa, I., & Funatogawa, T. (2019). *Longitudinal Data Analysis. Autoregresive Linear Mixed Effects Models*. Springer.

Garson, G.D. (2020). *Multilevel Modeling: Applications in STATA®, IBM® SPSS®, SAS®, R, & HLM™* . Sage Publications.

Grissom, R., & Kim, J. (2012). Effect size for research: Univariate and multivariate applications. 2nd edition. Routledge.

Grund, S., Lüdtke, O., & Robitzsch, A. (2019). Missing data in Multilevel Research. In S.E. Humphrey & J.M. LeBreton (Editors-in-Chief), *The Handbook of Multilevel Theory, Measurement and Analysis* (pp. 365-386). American Psychological Association.

Heck, R., & Thomas, S.L. (2020). *An Introduction to Multilevel Modeling Techniques. MLM and SEM Approaches Using Mplus*. Routledge.

Leke, C. A., & Marwala, T. (2019*). Deep Learning and Missing Data in Engineering Systems (Studies in Big Data Book 48)*. Springer.

Little, R.J.A., & Rubin D.B. (2020). *Statistical Analysis with Missing Data* (3rd edition). Wiley-Blackwell.

Livacic-Rojas P., Vallejo G., & Fernández, P. (2013). Covariance structures selection and Type I Error rates in Split Plot designs. *Methodology. European Journal of Research Methods for the Behavioural and Social Sciences, 9*(4), 129-136. https://doi.org/10.1027/1614-2241/a000058.

Livacic-Rojas P., Vallejo G., & Fernández, P. (2017). Power of modified Brown-Forsythe and Mixed-Model Approaches in Split-Plot Designs. Methodology. *European Journal of Research Methods for the Behavioural and Social Sciences, 13*(1), 9-22. https://doi.org/10.1027/1614-2241/a000124

Liu, X. (2016). *Methods and Applications of Longitudinal Data Analysis*. Academic Press.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*, 156-166. https://doi.org/10.1037/0033-2909.105.1.156

Molenberghs, G., Fitzmaurice, G., Kenward, M., Tsiatis, A., & Verbeke, G. (2015). *Handbook of Missing Data Methodology*. CRC Press.

Raghunathan, T. (2016). *Missing Data Analysis in Practice*. CRC Press.

Raghunathan, T., Berglund, P., & Solenberger, P. (2018). *Multiple Imputation in Practice*. CRC Press.

Ripley, B. D. (1987). *Statistical Inference for Spatial Processes*. Cambrigde University Press.

Sullivan, T.R., Yelland L.N., Lee K.L., Ryan, P., & Salter, A.B. (2017). Treatment of missing data follow-up studies of randomized controlled trials: A systematic review of the literature. *Clinical Trials, 14*(4), 387-395. https://doi.org/10.1177/1740774517703319

Vale, C. D., & Maurelli, V. A. (1983). Simulating multivariate non normal distributions. *Psychometrika*, *48,* 465-471. https://doi.org/10.1007/BF02293687.

Vallejo, G., Fernández, M. P., Livacic-Rojas, P. E., & Tuero-Herrero, E. (2011a). Selecting the best unbalanced repeated measures model. *Behavior Research Methods, 43*, 18-36. https://doi.org/10.3758/s13428-010-0040-1_

Vallejo, G., Fernández, P., Livacic-Rojas, P., & Tuero-Herrero, E.(2011b). Comparison of modern methods for analyzing repeated measures data with missing values. *Multivariate Behavioral Research, 46*, 900-937. https://doi.org/10.1080/00273171.2011.625320

Vallejo G., Tuero-Herrero E., Núñez J.C., & Rosário P. (2014). Performance evaluation of recent information criteria for selecting multilevel models in Behavioral and Social Sciences. *International Journal of Clinical and Health Psychology*, *14,* 48-57. https://doi.org/10.1016/S1697-2600(14)70036-5

Vallejo, G., Ato, M., Fernández, M.P., Livacic-Rojas, P.E., & Tuero-Herrero, E. (2016). Power analysis to detect treatment effect in longitudinal studies with heterogeneous errors and incomplete data. *Psicothema, 28*(3), 330-339. https://doi.org/10.7334/psicothema2016.92

Vallejo, G., Fernández, M. P., & Livacic-Rojas, P. E. (2018). Data analysis of incomplete repeated measures using a multivariate extension of the Brown-Forsythe procedure. *Psicothema, 30*(4), 434-441. https://doi.org/10.7334/psicothema2018

Vallejo, G., Fernández, M. P., & Livacic-Rojas, P. E. (2019). Sample size estimation for heterogeneous growth curve models with attrition. *Behavior Research Methods, 51*(3), 1216-1243. https://doi.org/10.3758/s13428-018-1059-y

Wilcox, R. (2017). *Introduction to Robust Estimation & Hypothesis Testing* (4rd ed.). Elsevier.

Yoo, J.E. (2013). *Multiple Imputation with Structural Equation Modeling. Using auxiliary variables when data are missing*. Academic Publishing.

Zhang, Z., & Yuan, K. (2018). *Practical Statistical Power Analysis. Using Web Power and R*. Isdsa Press.

Zhou, X.H., Zhou, Ch., Liu, D., & Ding, X. (2014). *Applied Missing Data Analysis in the Health Sciences*. Wiley & Sons.