



Confirmatory factor analysis with missing data in a small sample: cognitive reserve in people with Down Syndrome

Cristina Cañete-Massé^{1,3} · María Carbó-Carreté^{2,3} ·
María Dolores Figueroa-Jiménez⁵ · Guillermo R. Oviedo^{6,7} ·
Myriam Guerra-Balic⁶ · Casimiro Javierre⁸ · Maribel Peró-Cebollero^{1,3,4} ·
Joan Guàrdia-Olmos^{1,3,4}

Accepted: 11 October 2021 / Published online: 1 November 2021
© The Author(s) 2021

Abstract

The presence of missing data and small sample sizes are very common in social and health sciences. Concurrently to present a methodology to solve the small sample size and missing data, we aim to present a definition of Cognitive Reserve for people with Down Syndrome. This population has become an appealing focus to study this concept because of the high incidence of dementia. The accidental sample comprised 35 persons with DS (16–35 years). A total of 12 variables were acquired, four of them had missing data. Two types of multiple imputation were made. Confirmatory factor analysis with Bayesian estimations was performed on the final database with non-informative priors. However, to solve the sample size problem, two additional corrections were made: first, we followed the Jiang and Yuan (2017) schema, and second, we made a Jackknife correlation correction. The estimations of the confirmatory factor analysis, as well as the global fit, are adequate. As an applied perspective, the acceptable fit of our model suggests the possibility of operationalizing the latent factor Cognitive Reserve in a simple way to measure it in the Down Syndrome population.

Keywords Missing data · Small sample · Bayesian structural equation models · Down syndrome · Cognitive reserve

1 Introduction

In the field of social and health sciences the presence of sample anomalies is very common, specifically, samples are usually characterized by reduced sample size and missing data (Cohen 1990; Silvia et al. 2014). Obviously, these types of samples directly affect the validity of the data, the generalizing possibility of the results and the investigation quality. Moreover, most data analysis procedures are not designed for these abnormalities (Schafer and Graham 2002). However, in recent years Bayesian estimations have become popular

✉ Cristina Cañete-Massé
cristinacanete@ub.edu

Extended author information available on the last page of the article

in many scientific fields (Muthén and Asparouhov 2012; McNeish 2016; König and van der Schoot 2018, Smid et al. 2020), mainly because of the availability of popular softwares and because it adds some advantages over frequentists estimations. Regarding these advantages, one of the most interesting features is its flexibility to work with small samples sizes (MacNab et al. 2011; Smid et al. 2020).

In relation to missing data, multiple imputation (MI) techniques have become popular to replace the lost values with simulated data that allow the full cases analysis. These techniques are not exempt from controversy since they assume an arbitrary alteration of the original distributions and can modify some of their statistics (He 2010). There are a multitude of imputation possibilities and easy access, which implies an evident risk of erroneous selection of these techniques (Graham et al. 2006; Rhemtulla and Little 2012; Little et al. 2014; Smid et al. 2020).

Missing data is very frequent in studies with persons who have limitations to participate directly from themselves. Concurrently to present a methodology process to solve the small sample size and missing data, we aim to present a definition of Cognitive Reserve (CR) for people with Down Syndrome (DS). CR is the brain's ability to handle brain damage or disease, maintaining a stable level of function (Stern 2002, 2006, 2011, 2018). From our point of view, the study of CR concept in DS population would be of great advance to make a step forward in the dementia field. The prevalence and risk for dementia in this population is significantly different from general population or from people with Intellectual Disability (ID) of other aetiologies (Lott and Head 2019). In this population, as in general population, individual trajectories of cognitive changes are highly heterogenous, with some declining dramatically rapidly and others declining very slowly or even improving (Bull 2020; Temple et al. 2001). Given the variability that exists in DS, the implications of CR for individuals with DS deserve special attention to generate a more systematic definition (McGlinchey et al. 2014).

CR is considered to offer neuroprotective benefits against brain injury and disease by facilitating greater neural efficiency and/or enhanced capacity to develop compensatory cognitive mechanisms subsequent to the neuropathology onset (Jones et al. 2011; Stern 2002, 2006, 2011, 2018). Although this concept has been deeply studied, there are still measurement problems in order to operationalize CR by using different proxies. For instance, years of formal education is the most widely studied variable, and it is commonly used as the single proxy for the CR concept (Valenzuela and Sachdev 2006). However, CR could be also conceived as an hypothetical factor, not directly quantified by a unique measure, but as a latent construct. Therefore, nowadays CR is studied from a latent variable data analysis approach with more proxies as leisure activities (Ruiz-Contreras et al. 2012; Suchy et al., 2011), early family environment and genetic factors (Lee 2007), intelligence quotient (IQ; Stern 2006; Tucker-Drob et al. 2009; Tuokko et al. 2003); parent socioeconomic status (Richards and Sacker 2003); education attainment (Foubert-Samier et al. 2012) and physical activities (Gow et al. 2012).

To address the CR concept we have found some challenges related to the sample size and missing data, that we have had to solve through a suitable methodological procedure. The goal to define a concept of CR in DS population are included in a broadly study which the techniques used (e.g. aerobic capacity test) are difficult to administer in people with DS. Due to the complexity of the register techniques used, the sample obtained has been small and with enough missing data that could hinder the main study objectives.

Thus, the purposes of this paper are twofold; on the one hand, establish an heuristic for the treatment of missing data and small samples sizes, on the other, develop a measurement

model for the operational definition for CR in a DS population. As far as we know, there have been no published works in which a chained heuristic is proposed to solve both issues.

2 Method

2.1 Participants

The sample was composed of 35 persons with DS between 16 and 35 years old ($M=24.4$ and $SD=5.42$); 25.7% were women. The sampling was accidental, and recruitment was performed through several DS care institutions in Mexico (54.3%) and Spain (45.7%). The following inclusion criteria were applied: (a) age between 16 and 35 years old and (b) formal diagnosis of DS through karyotyping. Regarding the exclusion criteria: (a) evidence of other comorbid diagnoses implying cognitive dysfunction; (b) impossibility of obtaining consent from legal tutors; and (c) presence of medication affecting cognitive functions. In order to assess the presence of other comorbid diagnoses implying cognitive dysfunction, the Dementia Screening Questionnaire for Individuals with Intellectual Disability (DSQIID) was used. It has a high internal consistency ($\alpha=0.91$; Deb et al., 2007). Thus, based on the mentioned criteria the final sample was composed of 31 participants.

The diagnostic of ID is based in the Spain government law about disability, which assigns a "handicap" percentage to every person with a disability to represent the condition severity. The disability percentage is assigned administratively based on all types of impairments (e.g., intellectual, physical, sensorial). In the present study, 58% of the sample had mild ID, 38.7% had moderate ID, 3.2% had severe ID. None of the participants had sever or profound ID. IQ was assessed using the Kaufman Brief Test of Intelligence (KBIT, Kaufman and Kaufman 1990). However, due to the no adaptation of the test for people with ID, it wasn't included in the CFA model. However, IQ had a mean in our sample ($n=21$) of 46.21 ($SD: 10.91$); with a minimum of 40 and a maximum of 80.

2.2 Instruments

These data are part of a larger protocol in which the relation between cerebral signal (*fMRI*) and variables connected with cognitive outcome, quality of life and physical activity is studied. In the current study, in addition of the age and ID diagnosis (*Personal Conditions*) described above, the variables assessed through the next scales and tests were included. For the *Quality of Life* assessment, the Spanish version of Personal Outcomes Scale was used (Carbó-Carreté et al. 2015). This scale aims to assess quality of life in people with ID on the basis of the eight domain quality of life model (Schalock and Verdugo 2002), which was arranged into three higher-order factors: Independence (personal development, self determination, interpersonal relations); Social participation (social inclusion, rights); Wellbeing (emotional wellbeing, physical wellbeing, material wellbeing) (Wang et al. 2010). The reliability is appropriate for the domains and, particularly, for the factors, with α values higher than 0.82.

The *Cognitive Outcomes* were assessed by three measurements tools: (a) the semantic fluency (number of words produced during 1 min in two conditions, things to buy in a supermarket and names of colors); (b) the phonological fluency (number of words produced beginning with the letter "p" during 1 min) and (c) administration of the Frontal Assessment Battery (FAB) with an internal consistency of $\alpha=0.78$ (Dubois et al. 2000).

In relation the *Physical Activity* there were four assessments: (a) the first was related to the level of practice of physical activity, through the Level of Physical Activity Scale (LPAS) (Carbó-Carreté et al. 2016). The internal consistency of this scale is $\alpha=0.79$. The next two measurements were about physical activity and sedentary level by using GT3x Actigraph accelerometers (ActiGraph, Fort Walton Beach, FL, USA) (Oviedo et al. 2017). Specifically, (b) the Total Physical Activity (Total PA) were calculated by assessing the accelerations of the device on the longitudinal axis 1 (counts/min) and (c) the Vector Magnitude (counts/min), understood as the device accelerations on the longitudinal, sagittal and frontal axis, was calculated by the square root of the sum of the squares of the three axes. Finally, (d) the Cardiorespiratory Fitness was determined by obtaining the peak oxygen consumption relative to body weigh (relative VO_2 peak; ml/kg/min).

2.3 Procedure

The applied protocol was approved by the Bioethical Committee of the Universitat de Barcelona (16/03/2017) and the data was registered between March 2018 and July 2019. Informed consent of parents or tutors in legal charge of every person with DS was obtained. Moreover, informed consent from the participants with DS was obtained. All the participants were evaluated in two register sessions by previously trained investigators. The administration sequence was the same for all the participants, and the scales referenced above were administered first to avoid fatigue bias.

For the quality of life and for the Level of Physical Scale parent's participation was accepted. For the Cognitive Outcomes and the rest of Physical Activity it was required the direct participation of the person with DS. The researchers who administered the scales and the physical tests were specific professionals trained to do this assessments. In the Mexican and Spanish administration, the same instruments were used in all the registers to guarantee the comparability of the data and avoid measurement biases.

2.4 Statistical analysis and heuristics

First, the observed distribution of the variables was studied to evaluate the impact of the missing data in every variable. Of the twelve variables studied, only 4 presented missings: ID level (11.4% of missings), Cardiorespiratory fitness (22.9% of missings), Total PA (25.7% of missings) and finally Vector Magnitude (25.7% of missings). Second, the acceptable maximum percentage for missing imputation was studied. To do so, the Little χ^2 test was applied to evaluate whether the missings followed a MCAR, which would be the more desirable pattern.

In the case of the categorical variable ID diagnostic, the prediction models based on multinomial logistic models and linear discriminant analyses were studied to identify a predictor variable combination that allowed the relevance probability to be applied to one of the four categories of ID diagnostic prediction. None of the models tested offered statistically significant results. This forced us to discard the prediction models with auxiliary variable prediction. Therefore, it was decided to use simple imputation based on linear interpolation.

For the quantitative variables, the imputations derived from the linear regression (LR) and predictive mean matching (PMM) models were obtained. In the case of LR models, the linearity between pairs of variables and the possible effects of collinearity were previously analyzed, and stepwise estimation models were chosen. For both cases, LR and PMM, a total of 10

simulated samples were obtained iterating from the initial solution, following Montenegro and Chesnut (2015), and each was analyzed to identify the best adjusted iteration.

With the aim of achieving a more exhaustive study of the imputation results, the mean SE of all imputed means was obtained from the Rubin (1976) estimation. This is a standard error (SE) estimation of the parameter mean of interest (a).

$$SE(\bar{a}) = \sqrt{\frac{1}{M} \sum_{k=1}^M S_k^2 + \left(1 + \frac{1}{M}\right) \left(\frac{1}{M-1}\right) \sum_{k=1}^M (a_k - \bar{a})^2}$$

where M is the number of generated databases, s_k is the SE on the K database, a_k is the parameter estimation on the k database, and \bar{a} is the estimated parameter mean and uses the factor $(1 + 1/M)$ that corrects the equation because the number of databases is finite.

However, the adjustment estimation proposed by Rubin (1976), does not make a direct comparison of the mean's SE obtained in the original distribution and in the solution after the imputation mechanism. To overcome these limitations, an adjustment index associated with the SE of the original and imputed distributions was defined. The basic expressions for calculating this index, called the imputation fit index (IFI), are shown below.

$$IFI_{mk} = |EE(\bar{x}_{m(o)}) - EE(\bar{x}'_{m(k)})| \cong 0$$

$$\overline{IFI}_m = \frac{1}{k} \left(\sum_{i=1}^k IFI_k \right)$$

$$S^2_{IFI_m} = \frac{\sum_{i=1}^k (IFI_{mk} - \overline{IFI}_m)^2}{k - 1}$$

$$Z_{IFIk} = \frac{IFI_{mk} - \overline{IFI}_m}{\sqrt{S^2_{IFI_m}}}$$

$$\left[\begin{array}{l} Z_{IFI} \leq -1.65 \\ \alpha = 0.05 \end{array} \right],$$

where m are the different variables and k are the different iterations, $EE(\bar{x}_{m(o)})$ is the mean SE of the initial variable (observed) and $EE(\bar{x}'_{m(k)})$ is the mean SE of the variable in the k imputation. By means of this index, the difference between the mean SE of an observed variable and the mean SE of the imputed variable is estimated. Obviously, depending on whether the result of this difference tends to 0, the two distributions are closer. However, when standardizing the IFI distribution, the fit values will be placed in the bottom tail of the distribution. In this way, $Z_{IFI} \leq -1.65$ was set as the reference value for the acceptance of H_0 , that is, the observed and imputed equality of distributions.

As the final part of the present work, the resulting database was submitted to CFA. Given the small sample size, Bayesian parameter estimations were made using Bayesian structural equation models (BSEM) with non-informative priors (Muthén and Asparouhov 2012; Asparouhov et al. 2015; Hoijtink and Van de Schoot 2017).

Obviously, it would be preferable to have informative priors (Smid et al. 2020), but we do not have solid empirical data to use them as informative priors as is the first time to our knowledge that CR in DS is studied. The correlation matrix used as the input was imputed using Jackknife correlation. This procedure consists of estimating Pearson’s correlations for each pair of variables, using as many samples as are obtained from removing one subject from the sample at a time. From the distribution of correlations thus obtained, the average correlation and its 95% confidence interval were estimated. To integrate the correlation matrix, the lower limit value was used for all cases to obtain correlations that were as attenuated as possible and, at the same time, that had the least possible bias. Parameter estimations were made using MPlus version 8.4 and ad hoc programming in R version 3.6.3.

Given the reduced sample size, it was not feasible to study the measurement adjustment model by the usual indicators. Therefore, Jiang and Yuan’s (2017) strategy was used. These authors, based on the proposals of Satorra and Bentler (1988) and Jung (2013), developed a fitting index design that is effective for multivariable normal distribution violation assumption and suitable for very small samples. Jing and Yuan’s proposal (2017) is based on four fitting index extractions: T_{COR1} , T_{COR2} , T_{COR3} and P_{COR4} . All of them oscillate between [0,1], and the right settings must be assumed in values of $T_{CORi} > 0.90$ and $P_{COR4} > 0.10$. They are defined as follows:

$$T_{COR1} = \frac{T_{ML}}{C_1}$$

where $c_1 = tr(\hat{U}\hat{\Gamma})/rank(\hat{U}\hat{\Gamma})$ is fulfilled by $rank < df$. This condition indicates that it is impossible to estimate more parameters in the CFA than those determined by the sample size, n should be lower than the number of non-duplicated units in the variances-covariances matrix (\hat{U}). Finally, ($\hat{\Gamma}$) is the reproduced matrix from the usual estimated parameters in the adjustment of χ^2 . Moreover, $T_{ML} = nF_{ML}(\hat{\theta})$, where $F_{ML}(\hat{\theta})$ is the usual discrepancy function $F_{ML}(\theta) = tr(S\Sigma^{-1}(\theta)) - \log|S\Sigma^{-1}(\theta)| - p$.

From the previous expressions, Satorra and Bentler (1988) defined TML as follows:

$$T_{ML} = \sum_{j=1}^{df} \tau_j \chi_{1j}^2$$

in which every parameter (τ_j) considers the statistical adjustment of χ_{1j}^2 . From this expression, two more indicators are proposed. The first is the T_{RML} indicator in a rescaled measure:

$$T_{RML} = \frac{T_{ML}}{r}$$

$$\text{being } r = \frac{tr(\hat{U}\hat{\Gamma})}{(p^* - q)}$$

where p^* is the number of different units in R and q is the number of parameters to estimate. A second indicator adjusted to the variable called T_{AML} is defined as follows:

$$T_{AML} = \frac{T_{ML}}{\alpha}$$

$$\text{being } \alpha = \frac{\text{tr}[(\hat{U}\hat{\Gamma})^2]}{\text{tr}(\hat{U}\hat{\Gamma})}$$

$$T_{COR2} = \frac{T_{ML}}{C_2}$$

where $c_2 = \frac{r+c_1}{2}$ with the condition that c_2 should be between r and c_1 and the value of T_{COR2} should be between the values of T_{COR1} and T_{RM}

$$T_{COR3} = \frac{\frac{T_{ML}}{r} + \frac{T_{ML}}{c_1}}{2}$$

$$p_{COR4} = \frac{(p_{RML} + p_{AML})}{2},$$

where p_{RML} and p_{AML} correspond to the significance levels associated with the statistics of the χ^2 distribution and with the degrees of freedom of the measurement model.

3 Results

3.1 Imputation estimation

The lost percentage effect was studied from the statistical estimation of Little’s χ^2 test to obtain a MCAR pattern. Little’s test indicates a reasonable adjustment to a characteristic MCAR pattern (Little’s χ^2 test = 34.024, $df=27$, $p > 0.05$).

In the case of the categorical variable ID diagnostic, the prediction models based on multinomial logistic models and linear discriminant analyses were studied but none of the models tested offered statistically significant results (for instance, in the case of discriminant analysis, the Wilks’ λ values ranged from 0.750 to 0.964).

In view of the impossibility of using multiple imputations in this variable, it was decided to use simple imputation based on linear interpolation. The results after applying simple imputation shown that the median, mode and absolute deviations median for both variables were the same.

In relation to the quantitative variables, the mean SE of all imputed means were calculated through Rubin (1976) estimation. Table 1 shows the SE results for every variable and for every imputation technique.

In all cases, the Rubin (1976) estimates are higher in the LR solution. Likewise, the distribution of Rubin’s proposal does not present a standardized solution that facilitates interpretation, nor does it fix or recommend reference values from which a suitable solution is considered in a given imputation. Therefore, the imputation fit index (IFI) were calculated. An iteration process was generated by fixing a maximum of iterations ($k=10$) with the PMM method. It is important to remember that, given the small sample, the number of iterations recommended (Schafer and Graham 2002) was doubled to guarantee that an adequate solution was obtained. After performing the different IFI estimations for each

Table 1 Imputed variables with multiple imputation techniques and SE calculated with Rubin's technique (1976)

Imputed variables	LR SE(\bar{a})	PMM SE(\bar{a})
Total PA	1618.889	703.483
Vector Magnitude	4263.350	46.121
Cardiorespiratory fitness	2.263	1.487

variable, values lower than or equal to -1.65 were found for the 3 variables. Table 2 shows the IFI results in the 10 iterations performed.

In this first step, an adequate value of IFI ($Z_{IFI} \leq -1,65$) was obtained from the imputations performed with 3 variables (Total PA, Vector Magnitude and cardiorespiratory fitness). In the third column of Table 2, the iterations number for an adequate solution are shown.

3.2 Confirmatory factor analysis

Once the initial data database was reconstructed, a measurement model was established to define the CR latent factor for people with DS. This model was based on the works mentioned in the introduction with some adaptations because it is addressed to a specific population. The model proposed is composed of four latent factors (personal conditions, quality of life, cognitive outcome and physical activity). Figure 1 shows these factors with their correspondent direct indicators.

As shown in the previous figure, we have four first order latent factors, each defined by the indicators in Fig. 1 and only one second-order factor. To guarantee the range condition described above, the maximum number of variables possible must be a number that allows the degrees of freedom to be equal or greater than 0. For this, the number of possible free parameters to be estimated must be known. In this case, using all the variables in Fig. 1, the free parameters would be the following: Matrix Λ = 12 factor saturations; Matrix Φ = 6 correlation coefficients (as many as correlations other than 1 among the first-order factors); Matrix Θ_δ = 12 estimates of the variance of the measurement error. There are no other parameters to estimate since no correlations have been established between the variances of the measurement residuals [$E(\theta_{\delta ij} \theta_{\delta km}) = 0$]. Therefore, the number of free parameters to estimate is 30. Given this value, the minimum number of possible variables is derived from the expression $df = \frac{1}{2} [(p \cdot p + 1)] - q$, where p is the possible number of variables and q is the number of parameters. In these terms, the number of variables (p) must be greater than 8 to guarantee that $df > 0$.

Table 2 Variables to impute, better result in Z_{IFI} and iteration in which the better result was found

Variables (m)	Z_{IFI}	Iteration (k)
Step 1: Iteration's start (max $k=10$)		
Total PA	- 2.07	2
Vector magnitude	- 2.06	2
Cardiorespiratory fitness	- 2.37	5

Step 1 indicates the first iteration round

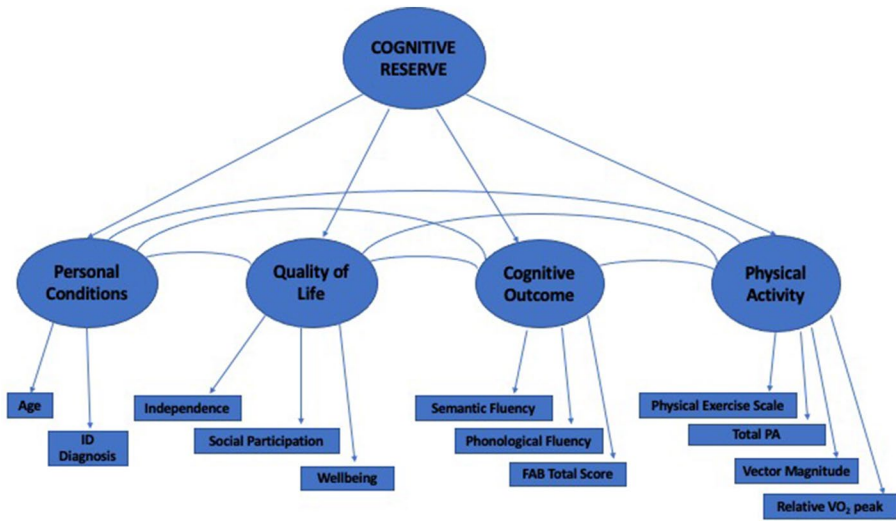


Fig. 1 Model Measurement of the Path Diagram Analysis

In our case, the condition is confirmed; therefore, the estimates are feasible using Bayesian estimation in accordance with what is proposed by Lee and Song (2004), Muthén and Asparouhov (2012) and Smid et al. (2020). Figure 2 shows the parameter estimation and all loading factors are statistically significant.

Regarding the correlations between factors, they are very low in all the cases. The higher value is between factor 2 and factor 3, that are, quality of life and cognitive outcome.

Finally, Table 3 shows the adjustment values of the proposed CFA model, showing both the usual indices and those proposed in the Jiang and Yuan's work (2017) as an alternative to the adjustments for small samples. The T_{CORR1} , T_{CORR2} and T_{CORR3} values are greater than 0.9, and P_{CORR4} is greater than 0.10. For all of these, we could say that the four corrections made by Jiang and Yuan (2017) demonstrate a good fit to the model. In the case of Bayesian Information Criteria (BIC) values, a lower value is preferable (Asparouhov et al. 2015).

4 Discussion

This study had to main objectives. The first aim of the present study was to propose a solution to overcome the problems of missing data and small sample size. MI was performed in the database, and two corrections for the small sample size were also done. The second aim of the paper was to estimate a CR model for DS people. A confirmatory factor analysis was developed and, as an applied point of view, the global fit of the model suggests the possibility of operationalizing the latent factor CR in a simple way to measure it in the Down Syndrome population.

More concretely, regarding the first aim of the study, all the variables followed MCAR distribution and two types of MI were made, LR and PMM. Once all variables used were completed, CFA with Bayesian estimations was performed. However, to solve the sample size problem, two additional corrections were made prior to the final analysis: Jiang and

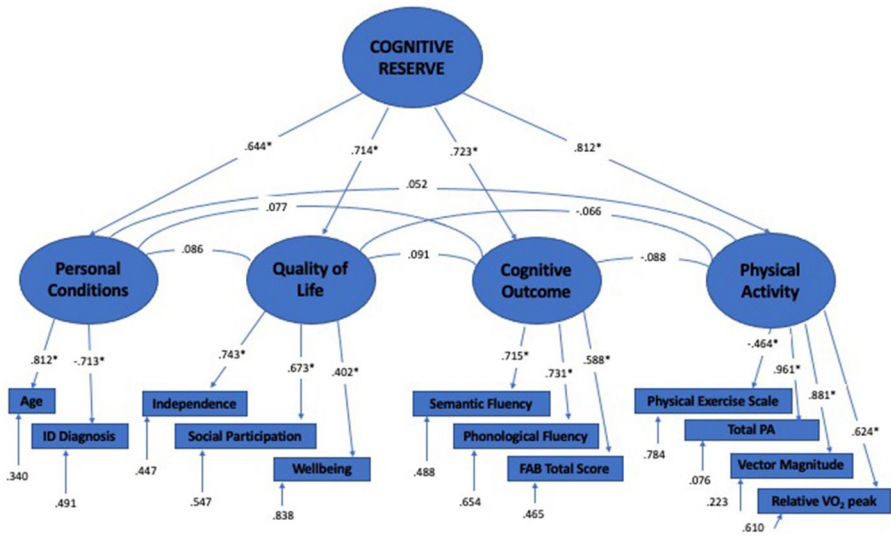


Fig. 2 Bayesian estimations of the free parameters (factor loading and correlation between factors) of the CFA with the model in Fig. 1. *Note:* correlations between factors are found in the first line, factor loadings (λ_i) in the second line and finally measurement error variance ($\theta_{\delta_{ii}}^2$) is found in the third line; *: $p < .001$

Yuan (2017) schema, and a Jackknife correlation correction. The estimations of the CFA, as well as the global fit, are adequate.

Obviously, we must assume that working with missing data and small sample size is not a good scenario. Even if we add some recommendations, it is not a perfect stage to work with. Therefore, what we aim to present in this paper is only valid when resampling is not viable and the population of interest is reduced. In the case of DS and considering the variables involved, this is the perfect situation to apply these techniques.

Past research on missing imputation has focused on which technique is the best or what point is acceptable to impute missing data (Rhetmulla and Little 2012; Little 2014). However, there is an important issue that directly affects these questions, the sample size. Our results focus on this aspect that affects all the decisions made from the first step. Our results illustrate several concerns about multiple imputation. Although the evaluation of imputations can be made by traditional estimations (Rubin 1976) and is widely used, this technique seems slightly weak because it does not directly compare the SE obtained in the original distribution and in the solution after multiple imputation. Moreover, the estimation does not evaluate which iteration works better. Therefore, to address these limitations, the use of our index (IFI) seems to be a good solution. This indicator allows us to know the point at which imputation generates a variable that behaves in a very similar way to the original one. Regarding the number of iterations made in the multiple imputations, it is important to note that although some authors argue that 5 iterations are sufficient to obtain good results (Montenegro and Chesnut 2015) and that the last iteration is normally the best, in the present paper we find that in many cases the best iteration is the second one. As mentioned above, the proposed IFI indicator has demonstrated to be a good tool to evaluate the quality of missing imputation.

In reference to the posterior analysis, a Bayesian CFA was conducted with non-informative priors. As noted, this analysis attempted to operationalize the concept of CR in DS.

Table 3 Statistical significance and values of the parameters of the model of measurement in Fig. 1

95% Confidence interval for the difference between the observed and the replicated Chi-square values	BIC	T_{COR1}	T_{COR2}	T_{COR3}	P_{COR4}
19.12 to 58.12	156.71	.971	.962	.933	.312

BIC: Bayesian Information Criteria

Regarding the present analysis, two specific points should be highlighted. First, we do not recommend the use of frequentist inferential techniques because of the limited sample size. Second, the use of Bayesian estimation could generate the problem of using informative or noninformative priors (Smid et al. 2020). These authors also note that the analysis with informative priors would be much stronger and much more potent in the case of small sample size, but many times, as in our case, this information is not available. The corrections that we proposed, based on Jiang and Yuan's (2017) seem to be a good approximation to study the global fit of BSEM. In the same line, our results show the capacity of conservative strategies such as the Jackknife correlation (in particular, the use of the low confidence interval value) to estimate the R matrix as an input to BSEM. The combination of both strategies makes it possible to generate a statistical framework that is more adequate to solve the small sample problem and limitations.

As an applied point of view, and regarding the second aim of this paper, the acceptable fit of the CFA model suggests the possibility of operationalizing the latent factors of CR to measure it in the DS population. The results of the individual Bayesian parameter estimation, as well the global fit, are evidence of this possibility. As proven above, Quality of Life, Cognitive Outcome, and Physical Activity can influence and mediate the concept of CR. Moreover, although years of formal education have been the principal variable studied to operationalize this concept (Valenzuela and Sachdev 2006), it is tough to measure in ID population. As mentioned above, the perspective of conceptualizing CR as a latent construct defined with different measures seems the clearer one. Obviously, according to the real properties of the sampling, this model does not allow any extrapolation for the general population and must be used strictly as an operational definition by our sample. However, it is an empirical confirmation and therefore an optimal measure for the operationalization of CR in persons with DS, at least in our sample. As it has been mentioned above, the prevalence of dementia in this population is much higher than in other ID populations and general population, this could be an initial phase in order to study the big differences encountered in DS population regarding the appearance of dementia.

These results are not exempt from some limitations. We did not study the statistical properties of the IFI to achieve a more adjusted significance or use simulations to estimate its sensitivity specificity. Another important limitation of this study is the use of noninformative priors in the Bayesian estimation of the CFA. Clearly, more efforts are needed to develop better-performing statistics with small sample sizes and missing data in the sample. However, in relation to the CR study, this paper provides the first approximation of this concept to the DS population and may provide a starting point to study the individual differences in the clinical and neuropathological appearance of dementias.

Authors contributions All authors contributed to the study conception and design. Cristina Cañete-Massé, María Carbó-Carreté, Maribel Peró-Cebollero and Joan Guàrdia Olmos made the first conceptualizations of the paper, investigation and methodology. María Carbó-Carreté, María Dolores Figueroa-Jiménez and Myriam Guerra-Balic provided the resources for the study. María Carbó-Carreté and Joan Guàrdia-Olmos administered the Project. Cristina Cañete-Massé and Joan Guàrdia-Olmos wrote the original draft preparation. Guillermo R. Oviedo made the validation of the project. Finally, supervision of the paper was made by María Carbó-Carreté, Maribel Peró-Cebollero and Joan Guàrdia-Olmos and finally, the final review of the paper was done by Guillermo R. Oviedo, Maribel Peró-Cebollero and María Carbó-Carreté.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This study has been funded by the Spanish Ministry of Science, Innovation and Universities (PGC2018-095829-B-I00).

Declarations

Conflicts of interest The authors state that there is no conflict of interest with the data presented in this study.

Availability of data and material The datasets generated during and/or analysed during the current study are not publicly available due to ethical reasons but are available from the corresponding author on reasonable request.

Ethics approval The applied protocol was approved by the Bioethical Committee of the Universitat de Barcelona (16/03/2017).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References









- Asparouhov, T., Muthén, B., Morin, A.J.S.: Bayesian structural equation modeling with cross-loadings and residual covariances comments on Stromeier et al. *J. Manag.* (2015). <https://doi.org/10.1177/0149206315591075>
- Bull, M.J.: Down syndrome. *N. Engl. J. Med.* (2020). <https://doi.org/10.1056/NEJMra1706537>
- Carbó-Carreté, M., Guàrdia-Olmos, J., Giné, C.: Psychometric properties of the Spanish version of the personal outcomes scale. *Int. J. of Clin. Health Psychol.* (2015). <https://doi.org/10.1016/j.ijchp.2015.04.002>
- Carbó-Carreté, M., Guàrdia-Olmos, J., Giné, C.: The physical activity support needs and strategies scale: its development and use. *Intellect. Dev. Disabil.* (2016). <https://doi.org/10.1352/1934-9556-54.1.32>
- Cohen, J.: Things I have learned (so far). *Am. Psychol.* (1990). <https://doi.org/10.1037/0003-066X.45.12.1304>
- Deb, S., Hare, M., Prior, L., Bhaumik, S.: Dementia screening questionnaire for individuals with Intellectual Disabilities (DSQIID). *Br J. Psychiatr. J. Mental Sci.* (2007). <https://doi.org/10.1192/bjp.bp.106.024984>
- Dubois, B., Slachevsky, A., Litvan, I., Pillon, B.: The FAB: a frontal assessment battery at bedside. *Neurol.* (2000). <https://doi.org/10.1212/WNL.55.11.1621>
- Foubert-Samier, A., Catheline, G., Amieva, H., Dilharreguy, B., Helmer, C., Allard, M., Dartigues, J.F.: Education, occupation, leisure activities, and brain reserve: a population-based study. *Neurobiol. Aging* (2012). <https://doi.org/10.1016/j.neurobiolaging.2010.09.023>
- Graham, J.W., Taylor, B.J., Olchowski, A.E.: Planned missing data designs in psychological research. *Psychol. Methods* (2006). <https://doi.org/10.1037/1082-989X.11.4.323>
- Gow, A.J., Bastin, M.E., Muñoz Maniega, S., Valdés Hernández, M.C., Morris, Z., Murray, C., Royle, A.N., Starr, J.M., Deary, I.J., Wardlaw, J.M.: Neuroprotective lifestyles and the aging brain: activity, atrophy, and white matter integrity. *Neurol.* (2012). <https://doi.org/10.1212/WNL.0b013e3182703fd2>
- He, Y.: Missing data analysis using multiple imputation: getting to the heart of the matter. *Circ.: Cardiovasc. Qual. Outcomes* (2010). <https://doi.org/10.1161/CIRCOUTCOMES.109.875658>
- Hojjink, H., van de Schoot, R.: Testing small variance priors using prior-posterior predictive *P*-values. *Psychol. Methods* (2017). <https://doi.org/10.1037/met0000131>
- Jiang, G., Yuan, K.H.: Four new corrected statistics for SEM with small samples and nonnormally distributed data. *Struct. Equ. Model.: Multidiscip. J.* (2017). 24(4): 479–494 <https://doi.org/10.1080/10705511.2016.1277726>
- Jones, R.N., Manly, J., Glymour, M.M., Rentz, D.M., Jefferson, A.L., Stern, Y.: Conceptual and measurement challenges in research on cognitive reserve. *J. Int. Neuropsychol. Soc.* (2011). <https://doi.org/10.1017/S1355617710001748>

- Jung, S.: Structural equation modeling with small sample sizes using two-stage ridge least-squares estimation. *Behav. Res. Methods* (2013). <https://doi.org/10.1080/10705511.2016.1277726>
- Kaufman, A. S., & Kaufman, N. L.: *Kaufman Brief Intelligence Test*. Circle Pines, MN: Am. Guidance Service (1990).
- König, C., van de Schoot, R.: Bayesian statistics in educational research: a look at the current state of affairs. *Educ. Res. Rev.* **70**(4), 486–509 (2018)
- Lee, S.Y., Song, X.Y.: Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivar. Behav. Res.* (2004). https://doi.org/10.1207/s15327906mbr3904_4
- Lee, H.: Understanding cognitive reserve through genetics and epidemiology. In: Stern, Y. (ed.) *Cognitive reserve. Theory and applications*, pp. 5–36. Taylor and Francis, New York, NY (2007)
- Little, T.D., Jorgensen, T.D., Lang, K.M., Moore, E.W.: On the joys of missing data. *J. Pediatr. Psychol.* (2014). <https://doi.org/10.1093/jpepsy/jst048>
- Lott, I.T., Head, E.: Dementia in Down syndrome: unique insights for Alzheimer disease research. *Nat. Rev. Neurol.* (2019). <https://doi.org/10.1038/s41582-018-0132-6>
- MacNab, Y.C., Malloy, D.C., Hadjistavropoulos, T., Sevigny, P.R., McCarthy, E.F., Murakami, M., Paholpak, S., Natarajan, S., Liu, P.L.: Idealism and relativism across cultures: a cross-cultural examination of physicians' responses on the Ethics Position Questionnaire (EPQ). *J. Cross-Cultural Psychol.* **42**(7), 1272–1278 (2011)
- McNeish, D.: On using Bayesian methods to address small sample problems. *Struct. Equ. Model.: Multidiscip. J.* (2016). <https://doi.org/10.1080/10705511.2016.1186549>
- McGlinchey, E., McCallion, P., McCarron, M.: Cognitive reserve: implications for research in dementia for people with down syndrome. *J. Appl. Res. Intellect. Disabil.* **27**(4) (2014)
- Montenegro, E.M., Oh, Y., Chesnut, S.: No le tema a los datos perdidos: enfoques modernos para el manejo de datos perdidos. *Actual. En Psicol.* (2015). <https://doi.org/10.15517/ap.v29i119.18812>
- Muthén, B., Asparouhov, T.: Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychol. Meth.* (2012). <https://doi.org/10.1037/a0026802>
- Oviedo, G., Travier, N., Guerra-Balic, M.: Sedentary and physical activity patterns in adults with intellectual disability. *Int. J. of Environ. Res. Public Health.* (2017). <https://doi.org/10.3390/ijerph14091027>
- Rhemtulla, M., Little, T.D.: Planned missing data designs for research in cognitive development. *J. of Cogn. and Devel.* (2012). <https://doi.org/10.1080/15248372.2012.717340>
- Richards, M., Sacker, A.: Lifetime antecedents of cognitive reserve. *J. Clin. Exp. Neuropsych.* (2003). <https://doi.org/10.1076/jcen.25.5.614.14581>
- Rubin, D.B.: Inference and missing data. *Biometrika* (1976). <https://doi.org/10.1093/biomet/63.3.581>
- Ruiz-Contreras, A.E., Soria-Rodríguez, G., Almeida-Rosas, G.A., García-Vaca, P.A., Delgado-Herrera, M., Méndez-Díaz, M., Prospéro-García, O.: Low diversity and low frequency of participation in leisure activities compromise working memory efficiency in young adults. *Acta Psychol.* (2012). <https://doi.org/10.1016/j.actpsy.2011.10.011>
- Satorra, A., Bentler, P. M. Scaling corrections for chi-square statistics in covariance structure analysis. In: *American Statistical Association 1988 Proceedings of Business and Economics Sections* (pp 308–313). Alexandria, VA: Am. Statistical Association (1988).
- Schafer, J.L., Graham, J.: WMissing data: our view of the state of the art. *Psychol. Methods.* (2002). <https://doi.org/10.1037/1082-989x.7.2.147>
- Schalock, R.L., Verdugo, M.A.: *Handbook on quality of life for human service practitioners*. American Association on Mental Retardation, Washington, DC (2002)
- Silvia, P.J., Kwapil, T.R., Walsh, M.A., Myin-Germeys, I.: Planned missing-data designs in experience-sampling research: Monte Carlo simulations of efficient designs for assessing within-person constructs. *Behav. Res. Methods* (2014). <https://doi.org/10.3758/s13428-013-0353-y>
- Smid, S.C., McNeish, D., Miočević, M., van de Schoot, R.: Bayesian versus frequentist estimation for structural equation models in small sample contexts: a systematic review. *Struc. Equ. Model.: multidiscip. J.* (2020). <https://doi.org/10.1080/10705511.2019.1577140>
- Stern, Y.: What is cognitive reserve? Theory and research application of the reserve concept. *J. Int. Neuropsychol. Soc.* (2002). <https://doi.org/10.1017/S1355617702813248>
- Stern, Y.: Cognitive reserve and Alzheimer disease. *Alzheimer Dis. Assoc. Disord.* (2006). <https://doi.org/10.1097/00002093-200607001-00010>
- Stern, Y.: Elaborating a hypothetical concept: Comments on the special series on cognitive reserve. *J. Int. Neuropsychol. Soc.* (2011). <https://doi.org/10.1017/S1355617711000579>
- Stern, Y., Arenaza-Urquijo, E.M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W.S., Okonkwo, O., Scarmeas, N., Soldan, A.,

- UdehMomoh, C., Valenzuela, M., Vemuri, P., Vuoksimaa, E.: Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's Dement.* (2018). <https://doi.org/10.1016/j.jalz.2018.07.219>
- Suchy, Y., Kraybill, M.L., Franchow, E.: Instrumental activities of daily living among community-dwelling older adults: Discrepancies between self-report and performance are mediated by cognitive reserve. *J. Clin. Experim. Neuropsych.* (2011). <https://doi.org/10.1080/13803395.2010.493148>
- Temple, V., Jozsvai, E., Konstantareas, M.M., Hewitt, T.A.: Alzheimer dementia in Down's syndrome: the relevance of cognitive ability. *J. Intellect. Disabil. Res.* (2001). <https://doi.org/10.1111/j.1365-2788.2001.00299.x>
- Tucker-Drob, E.M., Johnson, K.E., Jones, R.N.: The cognitive reserve hypothesis: a longitudinal examination of age-associated declines in reasoning and processing speed. *Dev. Psychol.* (2009). <https://doi.org/10.1037/a0014012>
- Tuokko, H., Garret, D.D., McDowell, I., Silverberg, N., Kristjansson, B.: Cognitive decline in high-functioning older adults: Reserve or ascertainment bias? *Aging Ment. Health* (2003). <https://doi.org/10.1080/1360786031000120750>
- Valenzuela, M.J., Sachdev, P.: Brain reserve and cognitive decline: a non parametric systematic review. *Psychol. Med.* (2006). <https://doi.org/10.1017/S0033291706007744>
- Wang, M., Schalock, R.L., Verdugo, M.A., Jenaro, C.: Examining the factor structure and hierarchical nature of the quality of life construct. *Am. J. Intellect. Dev. Disabil.* (2010). <https://doi.org/10.1352/1944-7558-115.3.218>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Cristina Cañete-Massé^{1,3}  · Maria Carbó-Carreté^{2,3}  ·
María Dolores Figueroa-Jiménez⁵  · Guillermo R. Oviedo^{6,7}  ·
Myriam Guerra-Balic⁶  · Casimiro Javierre⁸  · Maribel Peró-Cebollero^{1,3,4}  ·
Joan Guàrdia-Olmos^{1,3,4} 

¹ Department of Social Psychology and Quantitative Psychology, Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain

² Serra Hunter Fellow, Department of Cognition, Development and Educational Psychology, Faculty of Psychology, University of Barcelona, Barcelona, Spain

³ Institute of Neuroscience, Universitat de Barcelona, Barcelona, Spain

⁴ UB Institute of Complex Systems, Universitat de Barcelona, Barcelona, Spain

⁵ Department of Health Sciences, Centro Universitario de los Valles, Universidad de Guadalajara, Guadalajara, México

⁶ Faculty of Psychology, Education and Sport Sciences Blanquerna, Universitat Ramon Llull, Barcelona, Spain

⁷ School of Health Sciences Blanquerna, Universitat Ramon Llull, Barcelona, Spain

⁸ Department of Physiological Sciences, Universitat de Barcelona, Barcelona, Spain