**RESEARCH ARTICLE**

Statistics in Medicine WILEY

# Dynamic latent variable models for the analysis of cognitive abilities in the elderly population

## Silvia Bianconcini[1] | Silvia Cagnone[1]

[1]Department of Statistical Sciences, University of Bologna, Bologna, Italy

**Correspondence**
Silvia Cagnone, Department of Statistical Sciences, University of Bologna, via delle Belle Arti, Bologna 41 - 40126, Italy.
Email: silvia.cagnone@unibo.it

Cognitive functioning is a key indicator of overall individual health. Identifying factors related to cognitive status, especially in later life, is of major importance. We concentrate on the analysis of the temporal evolution of cognitive abilities in the elderly population. We propose to model the individual cognitive functioning as a multidimensional latent process that accounts also for the effects of individual-specific characteristics (gender, age, and years of education). The proposed model is specified within the generalized linear latent variable framework, and its efficient estimation is obtained using a recent approximation technique, called dimensionwise quadrature. It provides a fast and streamlined approximate inference for complex models, with better or no degradation in accuracy compared with standard techniques. The methodology is applied to the cognitive assessment data from the Health and Retirement Study combined with the Asset and Health Dynamic study in the years between 2006 and 2010. We evaluate the temporal relationship between two dimensions of cognitive functioning, that is, episodic memory and general mental status. We find a substantial influence of the former on the evolution of the latter, as well as evidence of severe consequences on both cognitive abilities among less-educated and older individuals.

**KEYWORDS**
generalized linear latent variable models, health mental status, intractable likelihood, vector autoregressive process

## 1 | INTRODUCTION

Cognitive functioning is a key indicator of overall individual health. Dimensions of cognitive skills are potentially important but often neglected determinants of the central economic outcomes that shape overall well-being over the life course. There exists enormous variation among households in their rates of wealth accumulation, their holdings of financial assets, and the relative risk in their chosen asset portfolios that have been proven difficult to explain by conventional demographic variables.[1]

   Identifying factors related to cognitive status and the maintenance of cognitive functioning, especially in later life, is of peculiar importance. Because of advances in public health and biomedicine, Americans are living longer than ever before. US national statistics indicate that the number of individuals over the age of 65 is expected to double in the next 35 years,[2] but living longer is not always an unmixed blessing. It is often associated with a greater prevalence of chronic conditions, including cognitive impairment that are largely age related.[3,4] Cognitive functioning is likely to impact one's

ability to work and play a role in retirement, particularly in the modern labor market which increasingly consists of jobs that require cognitive abilities and competence. These issues are likely to become much more relevant in light of changes to economic policies that are designed to encourage older workers to remain in the workforce until later ages. Hence, it is fundamental to understand which cognitive domains carry the most information on the earliest signs of cognitive impairment, and which subject characteristics are associated with a faster decline. This is not an easy task, being cognitive functioning a multidimensional construct composed by different attributes that operate simultaneously and differently according to individual-specific characteristics.

Longitudinal data of the type collected in household panel studies and population registers provide rich information on changes in individuals' behaviors and in their demographic and socioeconomic characteristics. However, cognitive measures are often not included in survey instruments because reliable assessments are too difficult and time-consuming to administer in a survey format. In this regard, the Health and Retirement Study combined with the Asset and Health Dynamic study (HRS/AHEAD) are among the best data sources from which to derive population-based estimates of cognitive impairment and to study a broad set of relevant outcome information. HRS/AHEAD is a nationally representative longitudinal study conducted by using telephone survey instrument on adults aged 50 or older in the United States that is sponsored by the National Institute of Aging and conducted by the University of Michigan every 2 years. The data are available at http://hrsonline.isr.umich.edu/. Several aspects of cognitive functioning have been assessed on more than 30 000 people (more than 17 000 families) during the past 25 years, from 1992 to 2016.

Promising approaches adopted in recent researches performed longitudinal analyses of the HRS/AHEAD cognitive measures to better describe age trends in cognition, and more generally, to identify determinants in the temporal pattern of the individual cognitive functioning among older adults.[5,6] Latent structures have been specified to determine common unobservable domains, that is, episodic memory and general mental status, underlying the cognitive tests in the HRS/AHEAD study, and growth models have been used to prove that cognitive functioning is associated with age,[5,7] gender,[8,9] and levels of education.[10,11] Recently, these data have been analyzed to identify homogeneous groups of respondents with similar cognitive profiles.[12,13]

Due to the decision to adopt telephone administration, HRS/AHEAD battery does not provide a comprehensive assessment of cognitive functioning, not including indicators of relevant cognitive domains, such as processing speed and executive function. However, although it has been found to measure only two cognitive constructs, it provides an acceptable set of items to assess dementia and, in particular, Alzheimer's disease (AD). AD is characterized by cognitive deficits that gradually affect overall cognitive functioning, but most studies consider episodic memory the most relevant, among the cognitive domains, in the preclinical phase of AD.[14-17] Indeed, the AD patients show deficits in episodic memory from the very early stages of the disease and hence this construct could constitute an important early marker to identify individuals at higher risk for developing AD.

The present article draws ideas from previous works and adds to their perspectives in terms of results as well as methodology. We propose to consider both aspects, the multifaceted nature of cognition and the longitudinal evolution of its components jointly rather than separately. A general framework that includes a large variety of latent variable models is represented by the generalized linear latent variable models (GLLVMs),[18] according to which the entire set of the responses given by an individual to a certain number of cognitive items, called the response pattern, is expressed as a function of one or more latent variables through a monotone differentiable link function. GLLVM for panel data model the associations of the latent and observed variables across time using random effects, latent variables, or both.[19-21]

We analyse the HRS/AHEAD longitudinal cognitive data through a generalization of the GLLVM for multidimensional and longitudinal data that is threefold. First of all, we test for the multidimensional factorial structure in the cognitive data at each time point, as well as for its measurement invariance over time. The interrelationship induced by repeatedly collecting multiple responses is accounted for by conditioning on the time-dependent latent cognitive domains and on item-specific random effects. Second, the structural part of the model describes nonstationary and general dynamics of multiple latent cognitive domains over time, and evaluates crosslagged effects among them. Finally, the impact of time-varying and time invariant individual characteristics on both episodic memory and mental status dynamics is accounted for.

A potential barrier to the application of latent variable models is the computational challenge presented by typically large datasets. National panel studies usually have several thousands of respondents which, when combined with multiple waves of measurement and a large choice set, renders existing (likelihood-based and Bayesian) estimation approaches unfeasible. If the observed cognitive items are of different nature, either continuous and discrete as in the HRS/AHEAD study, the estimation of these models is cumbersome. It can be carried out using a full information maximum likelihood method via either the EM algorithm or direct maximization, but, in both cases, the integrals involved in the likelihood

**TABLE 1** Descriptive statistics for the observed sample in wave 8 (2006)

| Variable | Details | Mean | SD | Min | Max |
| --- | --- | --- | --- | --- | --- |
| Gender | Male | 0.56 | – | – | – |
| Age | (in 2006) | 70.11 | 3.86 | 63 | 88 |
| Education | (in years) | 12.63 | 3.11 | 0 | 17 |

computation have no analytical solutions and need to be approximated. Classical quadrature techniques, such as the adaptive Gauss Hermite,[22,23] that represents the gold standard in the GLLVM framework, are already unfeasible when four items are analyzed at three different time points. Hence, these approximations are not applicable in our study since we need to consider more than four cognitive items to reliably measure the latent cognitive domains. Alternatively, the widely applied Laplace approximation[24,25] is known to provide inaccurate estimates in presence of discrete data.[26] To overcome these main limitations of standard techniques, we make use of a new approximation method, known as dimensionwise quadrature[27] that reduces the dimension of the multidimensional integrals, and makes the computation feasible also when the number of latent variables is large. The proposed approach provides a higher order approximation than the Laplace one but does not require any derivative computation, hence it is very simple to implement. Furthermore, for multidimensional and longitudinal data considered in this article, the corresponding estimators are asymptotically as accurate as the adaptive Gauss Hermite estimators.[27]

## 2 | A GENERAL MODEL FOR MULTIDIMENSIONAL LONGITUDINAL COGNITIVE DATA

### 2.1 | Preliminaries

The study focuses on HRS/AHEAD interviews of a subsample of 2941 spouses of the households' primary respondents, resulting after a listwise deletion of the missing data on all the cognitive variables of interest from wave 8 (2006) to wave 10 (2010). The sample is mainly composed by males (56%), the average age (at 2006) is equal to 70.11 years, and the average years of education are 12.63, as detailed in Table 1. These covariates have been widely recognized in the literature as predictors of the individual cognitive functioning.[6]

The cognitive measures in the HRS/AHEAD study used in this article include performance on (1) immediate (*IR*) and (2) delayed (*DR*) free recall of a list of 10 nouns (for a possible score of 0-10 on each measure), (3) serial 7s (*SER7s*), a working memory and mental processing task in which respondents counted backward from 100 by 7s for a total of five trials (for a possible score of 0-5), and mental status binary measures, including (4) naming the US president (*PRES*) and (5) vice president (*VCPRES*) by last name, (6) naming the cactus plant (*PLANT*) on the basis of a brief description, and (7) providing the year (*YEAR*) for an assessment of time orientation.

In previous studies on HRS/AHEAD data, these items have been converted in scores and treated as continuous. Here, we treat them as they are, that is as count and binary items. A descriptive analysis of their frequency tables at each wave allows us to get a first insight on the cognitive functioning pattern over time. Table 2 reports the percentages of response categories for the items *IR*, *DR*, and *SER7s* across the three waves. We can observe that for *IR* the highest percentages correspond to five and six immediately recalled words, and for *DR* to four and five words recalled after 5-min delay. For both items, the percentages of recalled words up to five tends to increase across time, whereas the reverse occurs in recalling more than five words, providing signs of potential cognitive impairment over time. For *SER7s* we can observe that 48.69% of individuals succeeded in all the five trials at wave 8, but the performance worsens over time.

Table 3 reports the percentages of correct responses for the remaining binary items. Differently from the previous results for count variables, the percentages of correct responses are very high and stable over time for all the mental status items, apart from *VCPRES* for which we observe a noticeable decrease of correct responses from wave 8 to wave 10.

The similar behavior of subsets of these cognitive measures gives insights on the presence of different underlying cognitive domains characterized by specific patterns over time. A generalized linear and latent variable approach would allow us to determine the contribution of each item to the latent cognitive variables and then to study their dynamics over time.

In recent studies,[5,6] a four-step procedure, based on the specification of different latent variable models at each stage, has been proposed to characterize the basic structure and age changes in these cognitive variables.

**TABLE 2**  Percentages of category responses for HRS/AHEAD items *IR*, *DR*, and *SER7s* in the three waves

| Items | Categories | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| *IR8* | 0.238 | 0.408 | 1.700 | 6.970 | 15.539 | 25.740 | 24.923 | 15.607 | 6.630 | 1.768 | 0.476 |
| *IR9* | 0.204 | 0.476 | 2.278 | 7.344 | 18.259 | 24.923 | 23.767 | 15.097 | 5.678 | 1.768 | 0.204 |
| *IR10* | 0.374 | 0.714 | 3.944 | 10.507 | 18.463 | 26.454 | 21.455 | 11.629 | 4.862 | 1.360 | 0.238 |
| *DR8* | 2.890 | 3.026 | 6.868 | 13.601 | 22.815 | 22.747 | 16.423 | 6.732 | 3.502 | 0.986 | 0.408 |
| *DR9* | 3.944 | 3.332 | 6.698 | 15.233 | 22.033 | 22.747 | 14.689 | 7.922 | 2.346 | 0.952 | 0.102 |
| *DR10* | 4.692 | 4.760 | 9.249 | 16.015 | 23.189 | 20.027 | 12.989 | 5.848 | 2.414 | 0.680 | 0.136 |
| *SER7s8* | 4.964 | 7.514 | 7.310 | 11.255 | 20.265 | 48.691 | – | – | – | – | – |
| *SER7s9* | 5.474 | 8.160 | 6.970 | 13.023 | 20.367 | 46.005 | – | – | – | – | – |
| *SER7s10* | 6.562 | 9.385 | 8.058 | 13.975 | 21.149 | 40.870 | – | – | – | – | – |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

**TABLE 3**  Percentages of correct responses for HRS/AHEAD items *YEAR*, *PLANT*, *VCPRES*, and *PRES* in the three waves

| Items | Waves | | |
|---|---|---|---|
| | 8 | 9 | 10 |
| *YEAR* | 98.810 | 98.334 | 96.974 |
| *PLANT* | 95.104 | 93.778 | 93.370 |
| *PRES* | 98.674 | 98.844 | 97.246 |
| *VCPRES* | 85.889 | 83.237 | 62.292 |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

At the first testing occasion, an exploratory factor analysis has been performed to define the factorial structure of the HRS/AHEAD cognitive measures, followed by a confirmatory factor analysis to define the cross-sectional relationships between the cognitive factors and a set of demographic variables. *IR* and *DR* have been found to be indicators of one factor expressing the *episodic memory* of individuals, that is the memory for newly acquired information, and the remaining items to be indicators of a second construct related to their *mental status*, a cognitive domain representing knowledge and use of established information. The proposed model is identified because of the so-called two-indicator rule for confirmatory factor models.[28]It is worth noting that in the HSR/AHEAD studies that applied factor models, *SER7s* has not be found to form a separate working memory factor as expected, probably because it is not the ideal measure for this component,[29] but it has resulted as an indicator of the general cognitive construct denoted with mental status.

At the third step, a latent multilevel model has been implemented in a longitudinal setting to evaluate the temporal factorial invariance of the common factors. These analyses have been applied to the factor scores obtained from the one and two-factor models under different invariance assumptions of the factorial structure. Finally, latent growth models have been fitted to observed measures and composite scores to analyze their growth/decline over time.

This four step procedure appears to have limitations related to the fact that the models were applied separately, and generally on different variables. To overcome these drawbacks, we propose a generalized linear and latent variable model that addresses all the issues raised by these authors simultaneously.

## 2.2 | Modeling the observed indicators: The measurement model

Let $\mathbf{y}_{it} = [y_{i1t} \ \ldots \ y_{ipt}]'$ denote $p$ cognitive measures observed for individual $i$ ($i = 1, \ldots, n$) at wave $t$ ($t = 1, \ldots, T$), $\mathbf{y}_i = [\mathbf{y}_{i1} \ \ldots \ \mathbf{y}_{iT}]'$ be the corresponding $pT$-dimensional vector, and let $\mathbf{x}_i = [x_{i1} \ \ldots \ x_{iH}]'$ represent the set of $H$ time-invariant covariates observed on the $i$th individual.

To properly measure the change in the individual cognitive condition, any proposed model has to account for the three sources of variability present in these data, that is (i) cross-sectional associations between the responses at a particular time point, (ii) crosslagged associations between different responses at different occasions, and (iii) the association between repeated measures of the same response over time. The first source of variability is accounted for the factorial structure underlying the responses $\mathbf{y}_{it}$ at each time point. It is based on $R$ time-dependent constructs $\mathbf{z}_{it} = [z_{it1} \ \ldots \ z_{itR}]'$ that represent the individual latent cognitive abilities at time $t$. Modeling the temporal evolution of $\mathbf{z}_i = [\mathbf{z}_{i1} \ \ldots \ \mathbf{z}_{iT}]'$ accounts for the crosslagged associations between different responses over time. Finally, to deal with the local dependence induced by the repeated measures on the same item over time, we further introduce a vector of item-specific random effects $\mathbf{u}_i = [u_{i1} \ \ldots \ u_{ip}]'$.[21] For the latter, alternative specifications have been discussed in the literature, based on the inclusion of an item-dependent autoregressive structure on the observations,[30] or on modeling the probability $g(\mathbf{y}_i \,|\, \mathbf{z}_i)$ directly via the inclusion of an extra-parameter of conditional dependence.[31,32]

Under the GLLVM framework for mixed data,[18] the probability associated to the individual response pattern $\mathbf{y}_i$ is expressed as follows

$$f(\mathbf{y}_i|\mathbf{x}_i) = \int_{\mathbb{R}^d} g(\mathbf{y}_i|\mathbf{z}_i, \mathbf{u}_i) h(\mathbf{z}_i, \mathbf{u}_i|\mathbf{x}_i) d\mathbf{z}_i d\mathbf{u}_i, \tag{1}$$

where the conditional distribution $g(\mathbf{y}_i|\mathbf{z}_i, \mathbf{u}_i)$ is usually referred to as the measurement part of the model and the distribution $h(\mathbf{z}_i, \mathbf{u}_i|\mathbf{x}_i)$ as the structural part. The dimension $d$ of the integral is equal to the total number of latent variables and random effects present in the model, that is $TR + p$.

**Assumption 1.** Given the time-dependent factors $\mathbf{z}_i$ and item-specific random effects $\mathbf{u}_i$, the observations $\mathbf{y}_i$ are independent, such that

$$g(\mathbf{y}_i|\mathbf{z}_i, \mathbf{u}_i) = \prod_{j=1}^p \prod_{t=1}^T g_j(y_{ijt}|\mathbf{z}_{it}, u_{ij}). \tag{2}$$

Each conditional distribution $g_j(y_{ijt}|\mathbf{z}_{it}, u_{ij})$ belongs to the exponential family and constitutes the random component of a generalized linear model. More specifically,

$$g_j(y_{ijt}|\mathbf{z}_{it}, u_{ij}) = \exp\left\{ \frac{y_{ijt}\theta_{ijt} - b_{ijt}(\theta_{ijt})}{\phi_j} + c_j(y_{ijt}, \phi_j) \right\},$$

where $\theta_{ijt}$ is the canonical parameter, $b_{ijt}(\theta_{ijt})$ and $c_j(y_{ijt}, \phi_j)$ are specific functions that assume different forms according to the different nature of $y_{ijt}$; $\phi_j$ is a scale parameter. For binary items, $g_j(y_{ijt}|\mathbf{z}_{it}, u_{ij})$ is the Bernoulli distribution of parameter $\mu_{ijt} = \pi_{ijt}(\mathbf{z}_{it}, u_{ij})$, and for count items $g_j(y_{ijt}|\mathbf{z}_{it}, u_{ij})$ we consider the Binomial distribution of parameters $\mu_{ijt} = \pi_{ijt}(\mathbf{z}_{it}, u_{ij})$ and $N_j$, the number of trials associated to the $j$th item, equal to 10 for $IR$ and $DR$, and to five for $SER7s$.

The systematic component of the model defines a linear predictor $\eta_{ijt}$ that depends on the latent variables and on the random effects as follows

$$\eta_{ijt} = \tau_{jt} + \boldsymbol{\alpha}'_{jt}\mathbf{z}_{it} + u_{ij}, \tag{3}$$

where $\boldsymbol{\alpha}_{jt}$ is the $R$-dimensional vector of the loadings associated to the latent variables $\mathbf{z}_{it}$ on $y_{ijt}$, $\tau_{jt}$ is the corresponding item- and time-dependent intercept, and $u_{ij}$ is the item-specific random effect. The link between the systematic component and the conditional mean of the random component is

$$\eta_{ijt} = \nu_j[\mu_{ijt}(\mathbf{z}_{it}, u_{ij})],$$

where $\nu_j$ can be any monotonic and differentiable function. The logit link function for both binomial and binary items is considered in this study.

Since the goal is to test substantive hypotheses about the common factors $\mathbf{z}_{it}$, we want to check if the content (meaning) of these factors is the same across occasions. The following assumption is then tested on the data.

**Assumption 2.** All the measurement parameters are invariant across occasions, that is

$$\tau_{jt} = \tau_j$$

and

$$\alpha_{jt} = \alpha_j,$$

for all $t$, $t = 1, \ldots, T$.

Under Assumption 2, the model is more parsimonious and avoids some possible identification problem that might arise with increasing the number of time points. However, especially in datasets with long follow-ups, some characteristics of the items could change over time due to possible birth cohort effects, ageing, differential loss, and other factors. Hence, Assumption 2 should always be tested in any specific dataset, rather than blindly imposed to the data.

To ensure identification of the model, one necessary condition is that all latent variables have a scale and an origin. Scales for the $\mathbf{z}_i$ factors can be provided either by fixing one loading per factor at a nonzero value at each occasion or by fixing the $R$ factor variances on one occasion at a nonzero value. On the other hand, origins for the $\mathbf{z}_i$ factors can be provided by either fixing $R$ intercepts, one for each factor, or by fixing the factor means at one occasion.[33] We fix the $R$ factor means at the first occasion equal to zero, and one loading for each latent factor equal to one at each occasion. In addition to these restrictions, we impose a simple structure on the loadings, such that each item is indicator of just one latent variable at each occasion.[34]

## 2.3 | A general dynamic model for the cognitive domains

The structural part of the model $h(\mathbf{z}_i, \mathbf{u}_i|\mathbf{x}_i)$ captures the longitudinal nature of the data, that is both the crosslagged and autocorrelations of the items over time. Different specifications for the moments of the joint density $h(\mathbf{z}_i, \mathbf{u}_i|\mathbf{x}_i)$ are related to different assumptions on the temporal dynamic of the data. Without loss of generality, we assume the latent variables and random effects to be normally distributed, that is

$$h(\mathbf{z}_i, \mathbf{u}_i|\mathbf{x}_i) \sim MVN\left(\boldsymbol{\mu} + \boldsymbol{\beta}'\mathbf{x}_i, \boldsymbol{\Sigma}\right),$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \ \ldots \ \boldsymbol{\beta}_T \ \mathbf{0} \ \ldots \ \mathbf{0}]'$ is a $(d \times H)$ matrix of regression coefficients for $\mathbf{x}_i$, being $\boldsymbol{\beta}_t, t = 1, \ldots, T$, a $R \times H$ matrix related to the latent domains $\mathbf{z}_{it}$ at time $t$ and $\mathbf{0}$ a $H$-dimensional null row vector associated to each random effect $u_{ij}, j = 1, \ldots, p$. $\boldsymbol{\mu} = [\boldsymbol{\mu}_\mathbf{z} \ \boldsymbol{\mu}_\mathbf{u}]'$ is the $d$-dimensional mean vector of the latent cognitive factors and random effects. Specifically, $\boldsymbol{\mu}_\mathbf{u}$ is the $p$-dimensional null vector, whereas $\boldsymbol{\mu}_\mathbf{z}$ is of dimension $TR \times 1$ and equal to $[\mathbf{0} \ \boldsymbol{\mu}_{\mathbf{z}_2} \ \ldots \ \boldsymbol{\mu}_{\mathbf{z}_T}]'$ due to identification constraints placed on the latent mean vector at the first occasion. Given the independence of the latent cognitive domains $\mathbf{z}_i$ and the item-specific random effects $\mathbf{u}_i$, $\boldsymbol{\Sigma}$ is a block diagonal matrix of the form

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_\mathbf{z} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_\mathbf{u} \end{bmatrix},$$

being $\boldsymbol{\Sigma}_\mathbf{z}$ a $TR \times TR$ full symmetric matrix, and $\boldsymbol{\Sigma}_\mathbf{u}$ a $p \times p$ matrix related to the item-specific random effects. Given that the associations among different items are all caught by the latent cognitive factors $\mathbf{z}_i$, $\boldsymbol{\Sigma}_\mathbf{u}$ is a diagonal matrix with generic element $\sigma^2_{u_j}, j = 1, \ldots, p$. Several assumptions can be made on the dynamics of the latent variables over time that in turn implies several forms of the matrix $\boldsymbol{\Sigma}_\mathbf{z}$. The values of these parameters could be unconstrained or depend further on model specification. For example, it is logical to expect that different cognitive domains are more strongly correlated when they are measured at closer time points. The choice of a specific structure can be driven by the data, main research questions, or by the chosen estimation method.[35]

Under the unstructured formulation for $\boldsymbol{\Sigma}_\mathbf{z}$, no assumptions are made on the temporal relationship of the latent variables. It follows that

$$\boldsymbol{\Sigma}_\mathbf{z} = \begin{bmatrix} \boldsymbol{\Sigma}_{1,1} & & & \\ \boldsymbol{\Sigma}_{2,1} & \boldsymbol{\Sigma}_{2,2} & & \\ \vdots & \vdots & \ldots & \\ \boldsymbol{\Sigma}_{T,1} & \boldsymbol{\Sigma}_{T,2} & \ldots & \boldsymbol{\Sigma}_{T,T} \end{bmatrix}, \tag{4}$$

where $\Sigma_{t,t} = \begin{bmatrix} \sigma^2_{z_{t1}} & & & \\ \sigma_{z_{t2},z_{t1}} & \sigma^2_{z_{t2}} & & \\ \vdots & \vdots & \vdots & \\ \sigma_{z_{tR},z_{t1}} & \sigma_{z_{tR},z_{t2}} & \cdots & \sigma^2_{z_{tR}} \end{bmatrix}$ and $\Sigma_{t,t'} = \begin{bmatrix} \sigma_{z_{t1},z_{t'1}} & & & \\ \sigma_{z_{t2},z_{t'1}} & \sigma_{z_{t2},z_{t'2}} & & \\ \vdots & \vdots & \vdots & \\ \sigma_{z_{tR},z_{t'1}} & \sigma_{z_{tR},z_{t'2}} & \cdots & \sigma_{z_{tR},z_{t'R}} \end{bmatrix}$, for $t, t'$ equal to 1, 2, ..., T, with $t \neq t'$.

$\Sigma_z$ is a full symmetric matrix with $\frac{TR(TR+1)}{2}$ parameters to be estimated. The elements on the main diagonal provide information on the variability of each cognitive domain over time. On the other hand, off-diagonal elements of $\Sigma_z$ give insights on the dependence of pairs of cognitive domains at the same and different occasions. A deep inspection of the estimates of these unconstrained parameters can provide justifications for more detailed model specifications, especially when covariances reduce as the time lag between the latent variables increases.

### 2.3.1 | Assessing the time-persistence and crosslagged effects of the different cognitive abilities

A specification for $\Sigma_z$ that takes the time ordering explicitly into account is the first-order vector autoregressive structure. This formulation presents the structural part of the model as a sequence of conditional distributions rather than a joint distribution with a completely free covariance matrix $\Sigma_z$. It expresses the dynamic nature of the latent variables and accounts for serial correlations by assuming that the latent variables at wave $t$, say, are related to those at previous waves only via the latent vector at wave $t-1$. We make use of the following vector autoregressive of order one process to model the evolution of the latent cognitive domains

$$\mathbf{z}_{i1} = \boldsymbol{\beta}'_1 \mathbf{x}_i + \boldsymbol{\varepsilon}_{i1} \tag{5}$$

$$\mathbf{z}_{it} = \boldsymbol{\mu}_{z_t} + \boldsymbol{\Phi}_{t,t-1} \mathbf{z}_{i,t-1} + \boldsymbol{\beta}'_t \mathbf{x}_i + \boldsymbol{\varepsilon}_{it} \qquad t = 2, \ldots, T, \tag{6}$$

where $\boldsymbol{\varepsilon}_{i1} \sim MNV(\mathbf{0}, \Sigma_{\varepsilon_1})$, being $\Sigma_{\varepsilon_1}$ a full and symmetric $R \times R$ matrix. $\boldsymbol{\Phi}_{t,t-1}$ is a full, nonsymmetric $R \times R$ matrix of autoregressive and crosslagged coefficients. We assume $\boldsymbol{\Phi}_{t,t-1} = \boldsymbol{\Phi}_1$ for $t = 2, \ldots, T$, such that the autoregressive parameters can be interpreted as measures of inertia or time-persistence of each cognitive domain. The elements on the main diagonal $\phi_{k,k}, k = 1, \ldots, R$, measure the autoregressive effect of the same latent construct over time. These parameters are related to the time it takes the individual to recover from a perturbation in that specific cognitive domain and restore a personal equilibrium on it. When the value of these parameters is close to zero, there is little carryover from one measurement occasion to the next. On the other hand, greater values of these coefficients indicate that there is considerable carryover between consecutive occasions, such that perturbations continue to have an effect on the cognitive domain under consideration on subsequent occasions.

The off-diagonal elements in $\boldsymbol{\Phi}_1$ evaluate the effects of one common factor on the other ones. Significant estimates of these parameters could be related to the differential predictive features of the cognitive domains on dementia,[16,17] in such a way that prevent decline on one dimension indirectly affects the evolution of the other dimension. Since it is reasonable to assume that it takes some time for a cognitive domain to affect the others, it is more appropriate to model lagged effects through $\boldsymbol{\Phi}_1$ instead of synchronized effects via the specification of $\boldsymbol{\Phi}_{t,t}$. For these reasons, $\boldsymbol{\varepsilon}_{it}$ is assumed to follow a multivariate normal density with $E(\boldsymbol{\varepsilon}_{it}) = \mathbf{0}$ and full symmetric covariance matrix $\Sigma_\varepsilon$, assumed to be time-invariant. Coherently with the previous studies on cognitive functioning, we are also interested in evaluating the influence of time-invariant covariates, such as age (at the baseline), gender, and years of education, on the latent cognitive domains. The vector $\mathbf{x}_i$ is constant over time, and the $R \times H$ matrix $\boldsymbol{\beta}_t$ contains the corresponding regression coefficients, assumed to be time-invariant, such that $\boldsymbol{\beta}_t = \boldsymbol{\beta}$, for $t = 2, \ldots, T$.

It is important to notice that, differently from common studies on cognitive functioning that analyzed the overall trends of a specific cognitive domain, the vector autoregressive model, specified in Equations (5) and (6), considers the dynamic of the multidimensional cognitive process as characterized by temporal changes that are not directly function of time but function of synchronized and crosslagged effects in the temporal evolution of its components. $\mathbf{z}_i$ has a multivariate normal distribution, with unconstrained $TR$-dimensional mean vector depending on $\boldsymbol{\mu}_{z_2}, \ldots, \boldsymbol{\mu}_{z_T}$, and $\boldsymbol{\beta}$, and implied covariance matrix $\Sigma_z$ being a constrained function of the parameters $\boldsymbol{\Phi}_1, \Sigma_\varepsilon, \Sigma_{\varepsilon_1}$ with the following form

$$\Sigma_z = \begin{bmatrix} \boldsymbol{\Gamma}_1(0) & \boldsymbol{\Gamma}_2(1)' & \cdots & \boldsymbol{\Gamma}_T(T-1)' \\ \boldsymbol{\Gamma}_2(1) & \boldsymbol{\Gamma}_2(0) & \cdots & \boldsymbol{\Gamma}_T(T-2)' \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{\Gamma}_T(T-1) & \boldsymbol{\Gamma}_T(T-2) & \cdots & \boldsymbol{\Gamma}_T(0) \end{bmatrix}.$$

Each $\boldsymbol{\Gamma}_t(l) = E(\mathbf{z}_t\mathbf{z}'_{t-l})$, $l = 0, \ldots, T-1, t = 1, \ldots, T$, is a $R \times R$ full and not symmetric matrix except for $l = 0$, being $\boldsymbol{\Gamma}_t(0)$ the concurrent covariance matrix at time $t$. $\boldsymbol{\Gamma}_t(l)$ can be expressed as a function of the parameters of the model as follows

$$\boldsymbol{\Gamma}_t(0) = \boldsymbol{\Phi}_1^{[t-1]}\boldsymbol{\Sigma}_{\varepsilon_1}\boldsymbol{\Phi}_1^{[t-1]'} + \sum_{i=0}^{[t-2]}\boldsymbol{\Phi}_1^i\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Phi}_1^{i'} \quad t = 1, \ldots, T.$$

$$\boldsymbol{\Gamma}_t(l) = \boldsymbol{\Phi}_1^{[t-1]}\boldsymbol{\Sigma}_{\varepsilon_1}\boldsymbol{\Phi}_1^{[t-l-1]'} + \sum_{i=0}^{[t-l-2]}\boldsymbol{\Phi}_1^{i+l}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\Phi}_1^{i'} \quad t = 2, \ldots, T; l = 1, \ldots, T-1.$$

A similar model specification has been recently developed in the Bayesian framework, but under the assumption of stationarity of the latent process.[32] It represents a special case of our model specification, where at the first wave $\mathbf{z}_1$ is treated as endogenous. That is, $\boldsymbol{\Sigma}_{\varepsilon_1}$ is set equal to $\boldsymbol{\Sigma}_\varepsilon$, and the mean vector at all occasions is assumed to be constant, $\boldsymbol{\mu}_{z_t} = \boldsymbol{\mu}_z$, and equal to the null vector, $\boldsymbol{\mu}_z = \mathbf{0}$ for all $t = 1, \ldots, T$. The estimation of this stationary autoregressive process requires all the eigenvalues of $\boldsymbol{\Phi}_1$ to be in modulus less than one, meaning that the process is stable. This latter condition is generally satisfied in presence of long follow-ups, as widely discussed in the time series literature.[36] On the other hand, in presence of short panel data, as in our case, the vector autoregressive process is generally nonstationary.

## 2.3.2 | Analysis of cognitive declines over time

An alternative specification to describe the (temporal) evolution of the latent cognitive domains would be through factor-specific random intercepts and slopes that affect $\mathbf{z}_i$ as in a standard multivariate growth model.[37] Coherently with previous researches on the HRS/AHEAD cognitive measures, this specification evaluates the cognitive decline of each cognitive domain, separately, by also accounting for the effect of specific determinants, such as age, gender, and years of education. However, differently from these previous studies, the specification of a multivariate growth model allows us to draw conclusions on the relationships among the random components of the different cognitive domains, that is on reciprocal influences on their initial status and decline over time.

The temporal dynamics of the each cognitive domain is described as follows

$$z_{itr} = \beta_{0ir} + \beta_{1ir}\Lambda_t + \varepsilon_{itr} \quad r = 1, \ldots, R; t = 1, \ldots, T; i = 1, \ldots, n,$$
$$\beta_{0ir} = \mu_{\beta_{0r}} + \boldsymbol{\gamma}_{\beta_{0r}}\mathbf{x}_i + \varsigma_{\beta_{0ir}}$$
$$\beta_{1ir} = \mu_{\beta_{1r}} + \boldsymbol{\gamma}_{\beta_{1r}}\mathbf{x}_i + \varsigma_{\beta_{1ir}}$$

For each factor, the latent curve can be assumed to be linear $[\Lambda_t = (t-1)]$ or nonlinear [some $\Lambda_t$'s freely estimated]. The corresponding growth components $\beta_{0ir}$ and $\beta_{1ir}$ are correlated and characterized by overall means equal to $\mu_{\beta_{0r}}$ and $\mu_{\beta_{1r}}$, variances $\sigma^2_{\beta_{0r}}$ and $\sigma^2_{\beta_{1r}}$, respectively, and covariance $\sigma_{\beta_{0r},\beta_{1r}}$. The effect of the time invariant covariates $\mathbf{x}_i$ on both the initial status and rate of change of each cognitive domain is assessed through the coefficients $\boldsymbol{\gamma}_{\beta_{0r}}$ and $\boldsymbol{\gamma}_{\beta_{1r}}$, respectively. The value of these parameters allows us to estimate average trajectories, for each latent factor, of different groups of individuals identified by each covariate.

The relationships among the growth components $\beta_{0ir}$, $\beta_{1ir}$, $\beta_{0ir'}$, and $\beta_{1ir'}$, $r, r' = 1, \ldots, R$ with $r \neq r'$, can be assessed in several ways.[37] The most general and easiest specification consists in the associative multivariate growth model, in which covariances among all the random components are estimated. Significant estimates suggest the existence of linear relationships between the initial status of two cognitive domains ($\sigma_{\beta_{0r},\beta_{0r'}}$), the initial status of one latent component and the rate of change of another one ($\sigma_{\beta_{0r},\beta_{1r'}}, \sigma_{\beta_{0r'},\beta_{1r}}$), and/or the rates of change (decline) of two latent variables ($\sigma_{\beta_{1r},\beta_{1r'}}$).

A deeper investigation of the covariance matrix of the random components can suggest more sophisticated specifications that include common factors to describe the relationship between the growth components of the latent cognitive domains, as in the "factor-of-curves".[38] In this model, second-order factors are specified to assess whether a higher order structure drives the relationship among the trajectories of all the latent cognitive domains as follows

$$\beta_{0ir} = \beta_{0i} + \boldsymbol{\gamma}_{\beta_{0r}}\mathbf{x}_i + \varsigma_{\beta_{0ir}} \quad r = 1, \ldots, R; i = 1, \ldots, n,$$
$$\beta_{1ir} = \beta_{1i} + \boldsymbol{\gamma}_{\beta_{1r}}\mathbf{x}_i + \varsigma_{\beta_{1ir}},$$

with $\begin{bmatrix} \varsigma_{\beta_{0ir}} \\ \varsigma_{\beta_{1ir}} \end{bmatrix} \sim MVN \left( \mathbf{0}, \begin{bmatrix} \sigma^2_{\beta_{0r}} & \sigma_{\beta_{0r},\beta_{1r}} \\ \sigma_{\beta_{0r},\beta_{1r}} & \sigma^2_{\beta_{1r}} \end{bmatrix} \right)$. The parameter $\beta_{0i}$ can be interpreted as individual cognitive functioning random intercept, and represents what is shared among all the cognitive domains at the initial time point, whereas $\beta_{1i}$ (ie, individual cognitive functioning random slope) is the shared growth pattern among the constructs over time (ie, they are growing, or changing over time, in a similar way). The expected value for the individual cognitive functioning random intercept, $\mu_{\beta_0}$, and slope, $\mu_{\beta_1}$, inform about the participants' average levels in the initial status and average rate of change, respectively, common to all the latent cognitive domains. On the other hand, their variances, $\sigma^2_{\beta_0}$ and $\sigma^2_{\beta_1}$, inform about the individual differences in these parameters.

## 3 | LIKELIHOOD ESTIMATION

The most commonly used approaches for fitting GLLVMs are maximum likelihood[18,21,25] and Markov chain Monte Carlo estimation.[20,32] In this study, we focus on the former. Recent computational advances have provided researchers with the tools to estimate models that incorporate longitudinal data with full information maximum likelihood. However, its feasibility when fitting latent variable models to multivariate longitudinal data is dependent upon practical concerns, as computational complexity greatly increases as the number of items (random effects) and latent variables rises.

Denoting with $\mathbf{b}_i = (\mathbf{z}_i, \mathbf{u}_i)$ the vector of latent variables and random effects, the log-likelihood function is given by

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(\mathbf{y}_i|\mathbf{x}_i, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log \int_{\mathbb{R}^d} g(\mathbf{y}_i|\mathbf{b}_i, \boldsymbol{\theta}_1) h(\mathbf{b}_i|\mathbf{x}_i, \boldsymbol{\theta}_2) d\mathbf{b}_i, \tag{7}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)'$ is the vector of the parameters to be estimated. Maximum likelihood estimation requires maximization of the integral with respect to $\boldsymbol{\theta}$. In presence of observed normal variables, the multidimensional integral can be evaluated analytically and maximization proceeds by standard methods such as EM algorithm, Fisher scoring or Newton-Raphson. However, in general, and in our specific case, integration is intractable.

Numerical quadrature-based methods represent a widespread solution to this problem and, among them, the adaptive Gauss Hermite quadrature is considered the gold standard.[23,39] Alternatively, the Laplace approximation avoids the integral computation and represents the easiest method to implement.[24] Both the Laplace approximation and the adaptive Gauss Hermite quadrature rely on the Taylor series expansion of the logarithm of the integrand in (7), denoted by $L(\mathbf{b}_i)$, around its mode $\mathbf{b}_{mo,i} = \arg\max_{\mathbf{b}_i \in \mathbb{R}^d} L(\mathbf{b}_i)$.[40] That is, by omitting the individual subscript $i$ for simplicity,

$$L(\mathbf{b}) = L(\mathbf{b}_{mo}) + \frac{1}{2}(\mathbf{b} - \mathbf{b}_{mo})'L^{(2)}(\mathbf{b}_{mo})(\mathbf{b} - \mathbf{b}_{mo}) + \sum_{k=3}^{\infty} \frac{1}{k!}[\otimes^{k-1}(\mathbf{b} - \mathbf{b}_{mo})']L^{(k)}(\mathbf{b}_{mo})(\mathbf{b} - \mathbf{b}_{mo})$$

$$= L(\mathbf{b}_{mo}) + \frac{1}{2}(\mathbf{b} - \mathbf{b}_{mo})'L^{(2)}(\mathbf{b}_{mo})(\mathbf{b} - \mathbf{b}_{mo}) + \nu(\mathbf{b}) \tag{8}$$

where

$$L^{(k)}(\mathbf{b}_{mo}) = \left. \frac{\partial \text{vec} L^{(k-1)}(\mathbf{b})}{\mathbf{b}} \right|_{\mathbf{b}=\mathbf{b}_{mo}}$$

and $\otimes^k(\mathbf{b} - \mathbf{b}_{mo})' = (\mathbf{b} - \mathbf{b}_{mo})' \otimes \ldots \otimes (\mathbf{b} - \mathbf{b}_{mo})'$, there being $k$ $(\mathbf{b} - \mathbf{b}_{mo})'$'s in the Kronecker product.[41]

Substituting the expansion (8) into the integral, the marginal probability associated to the individual response pattern results

$$f(\mathbf{y}|\mathbf{x}; \boldsymbol{\theta}) = (2\pi)^{d/2}|\boldsymbol{\Sigma}_{mo}|^{1/2}\exp\{L(\mathbf{b}_{mo})\}E_\phi\left[\exp\{\nu(\mathbf{b})\}\right] = f_L E_\phi\left[\exp\{\nu(\mathbf{b})\}\right], \tag{9}$$

where $f_L$ is the Laplace approximation of the integral, obtained by truncating the Taylor expansion up to the second order. The simplicity of the standard Laplace approximation has a cost related to the fact that the order of the approximation error cannot be improved, and it gets less adequate as the degree of discreteness increases.[26]

The inclusion in the Taylor series expansion of higher (than two) order terms $v(\mathbf{b})$ improves the accuracy of the estimates. The expected value $E_\phi\left[\exp\{v(\mathbf{b})\}\right]$ is computed with respect to the multivariate normal density function $\phi(\mathbf{b}; \mathbf{b}_{mo}, \boldsymbol{\Sigma}_{mo})$, whose mean vector is given by the mode $\mathbf{b}_{mo}$ and the covariance matrix is given by minus the inverse of the Hessian matrix of $L(\mathbf{b})$ evaluated at its mode, that is $\boldsymbol{\Sigma}_{mo}^{-1} = -L^{(2)}(\mathbf{b}_{mo})$.

$v(\mathbf{b})$ can be alternatively expressed as $L(\mathbf{b}) - L(\mathbf{b}_{mo}) - \frac{1}{2}(\mathbf{b} - \mathbf{b}_{mo})'L^{(2)}(\mathbf{b}_{mo})(\mathbf{b} - \mathbf{b}_{mo})$, such that the expected value in (9) results

$$E_\phi\left[\exp\{v(\mathbf{b})\}\right] = \frac{1}{f_L}\int_{\mathbb{R}^d}\frac{\exp[L(\mathbf{b})]}{\phi(\mathbf{b}; \mathbf{b}_{mo}, \boldsymbol{\Sigma}_{mo})}\phi(\mathbf{b}; \mathbf{b}_{mo}, \boldsymbol{\Sigma}_{mo})d\mathbf{b}. \tag{10}$$

Equation (10) resembles the classical transformation applied in the adaptive Gauss Hermite approximation to ensure the integrand to be sampled in a suitable range.[42] Rewriting it in terms of standardized variables, we obtain

$$E_\phi\left[\exp\{v(\mathbf{b})\}\right] = \frac{|\mathbf{C}_{mo}|}{f_L}\int_{\mathbb{R}^d}\frac{\exp[L(\mathbf{C}_{mo}\mathbf{b}^* + \mathbf{b}_{mo})]}{\phi(\mathbf{b}^*; \mathbf{0}, \mathbf{I})}\phi(\mathbf{b}^*; \mathbf{0}, \mathbf{I})d\mathbf{b}^*,$$

where $\mathbf{C}_{mo}$ is derived by the Cholesky decomposition of $\boldsymbol{\Sigma}_{mo}$. Hence, the Gauss Hermite approximation of the marginal density $f(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ is given by

$$f_{agh}(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = (2)^{d/2}|\mathbf{C}_{mo}|\sum_{q_1, \ldots, q_d} g(\mathbf{y}|\mathbf{b}^*_{q_1 \ldots q_d}, \boldsymbol{\theta}_1)h(\mathbf{b}^*_{q_1 \ldots q_d}, \boldsymbol{\theta}_2)w^*_{q_1} \ldots w^*_{q_d}, \tag{11}$$

where $\sum_{q_1, \ldots, q_d} = \sum_{q_1=1}^{n_q} \ldots \sum_{q_d=1}^{n_q}$, with $n_q$ be the number of quadrature points selected for each latent variable, $\mathbf{b}^*_{q_1 \ldots q_d} = \sqrt{2}\mathbf{C}_{mo}(b_{q_1} \ldots b_{q_d}) + \mathbf{b}_{mo}$, and $w^*_{q_m} = w_{q_m}\exp(b_{q_m}^2)$, $m = 1, \ldots, d$, being $b_{q_m}$ and $w_{q_m}$ the classical Gauss Hermite nodes and weights, respectively.

The adaptive Gauss Hermite quadrature provides more accurate estimates than the Laplace approximation, but it is computationally unfeasible with a large number of latent variables and random effects. It requires $n_q^d$ function evaluations at each iteration. As an example, if $d = 13$, that is we assume to observe seven items that measure two latent factors at each of three occasions, selecting only five quadrature points for each dimension ($n_q = 5$), it needs more than one million and two hundreds function evaluations at each iteration, being unfeasible with the most powerful computers available nowadays.

To improve the accuracy of the Laplace approximation by avoiding the computational burden of the adaptive Gauss Hermite, the expected value in (9) can be approximated using the following equivalent representation of the function $\exp\{v(\mathbf{b})\}$[27]

$$\exp\{v(\mathbf{b})\} = \sum_{w=0}^{d}\sum_{r=0}^{d-w}(-1)^w\binom{d-w}{r}\exp\{v_w(\mathbf{b})\}. \tag{12}$$

$\exp\{v_w(\mathbf{b})\} = \sum_{k_1=1}^{d-(w-1)}\sum_{k_2=k_1+1}^{d-(w-2)}\cdots\sum_{k_w=k_{w-1}+1}^{d}\exp\{v(\mathbf{b}_{k_1, \ldots, k_w})\}$, where $\mathbf{b}_{k_1, \ldots, k_w}$ is the vector $\mathbf{b}$ in which only the elements that occupy the positions $k_1, \ldots, k_w$ are free random variables, being the other $(d - w)$ fixed to the corresponding modes. $\sum_{1 \leq k_1 < \ldots < k_w \leq d} = \sum_{k_1=1}^{d-(w-1)}\sum_{k_2=k_1+1}^{d-(w-2)}\cdots\sum_{k_w=k_{w-1}+1}^{d}$ is the sum over all possible positions $k_1, \ldots, k_w$ in the vector $\mathbf{b}$.

The dimensionwise approximation is derived by truncating the sums that appear in Equation (12) to involve a smaller number, say $s$, of latent variables and random effects in the representation, with $s$ much smaller than $d$. That is,

$$\exp\{v(\mathbf{b})\} \approx \sum_{w=0}^{s}\sum_{r=0}^{s-w}(-1)^w\binom{d-w}{r}\exp\{v_w(\mathbf{b})\}$$

$$= \sum_{w=0}^{s}(-1)^{s-w}\binom{d-w-1}{s-w}\exp\{v_w(\mathbf{b})\}. \tag{13}$$

Substituting this approximation into the expected value (10), we obtain

$$
E_\phi\left[\exp\{\nu(\mathbf{b})\}\right] \approx \frac{1}{f_L}\sum_{w=0}^{s}(-1)^{s-w}\binom{d-w-1}{s-w}
$$

$$
\sum_{1\leq k_1<\cdots<k_w\leq d}\int_{\mathbb{R}^w}\frac{\exp[L(\mathbf{b}_{k_1\ldots k_w})]}{\phi(\mathbf{b}_{k_1\ldots k_w};\mathbf{b}_{mo},\boldsymbol{\Sigma}_{mo})}\phi(\mathbf{b}_{k_1\ldots k_w};\mathbf{b}_{mo},\boldsymbol{\Sigma}_{mo})d\mathbf{b}_{k_1\ldots k_w}. \tag{14}
$$

It is evident that the dimensionwise method consists in approximating the $d$-dimensional integral with the sum of one-, two-, up to $s$-dimensional integrals. Hence, by approximating the latter using classical Gauss Hermite nodes and weights, the dimensionwise approximation of the marginal density $f(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$ results

$$
f_{dim}(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}) = (2)^{d/2}|\mathbf{C}_{mo}|\sum_{w=0}^{s}(-1)^{s-w}\binom{d-w-1}{s-w}
$$

$$
\sum_{1\leq k_1<\cdots<k_w\leq d}\sum_{q_1,\ldots,q_w}g(\mathbf{y}|\mathbf{b}^*_{k_1,q_1;\cdots;k_w,q_w},\boldsymbol{\theta}_1)h(\mathbf{b}^*_{k_1,q_1;\cdots k_w,q_w},\boldsymbol{\theta}_2)w^*_{k_1,q_1}\cdots w^*_{k_w,q_w}, \tag{15}
$$

where $\sum_{q_1,\ldots,q_w}=\sum_{q_1=1}^{n_q}\cdots\sum_{q_w=1}^{n_q}$, with $n_q$ be the number of quadrature points selected for each latent variable, $\mathbf{b}^*_{k_1,q_1\ldots k_w,q_w}=\sqrt{2}\mathbf{C}_{mo}(b_{k_1,q_1}\ 0\ \ldots\ b_{k_w,q_w}\ 0)+\mathbf{b}_{mo}$, and $w^*_{k_m,q_m}=w_{q_m}\exp(b^2_{k_m,q_m})$, $m=1,\ldots,w$, being $b_{k_m,q_m}$ and $w_{q_m}$ the classical Gauss Hermite nodes and weights, respectively.

The approximation (15) is much less computational intensive than the adaptive Gauss Hermite method, being the number of function evaluations required at each iteration equal to $\sum_{w=0}^{s}\binom{d}{s-w}n_q^{s-w}$. For example, when $d=13$, if five quadrature points are selected for each dimension ($n_q=5$) and the 13-dimensional integral is approximated by the sum of unidimensional integrals ($s=1$), the number of function evaluations required at each iteration are 66. By including also bidimensional integrals ($s=2$) the number increases at 2016, and with three-dimensional integrals ($s=3$) it is a bit greater than 37 000 function evaluations.

For finite samples, the dimensionwise approximation achieves similar accuracy to the adaptive Gauss Hermite quadrature by selecting $s$ at maximum equal to three.[27] For multidimensional data, the two procedures share the same consistency rate asymptotically. Indeed, it is worth noticing that when $s=d$, that is no integral reduction is performed, we end up in the exact representation of $\exp\{\nu(\mathbf{b})\}$ (see Equation (12)), such as $f_{dim}(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$ equals the adaptive solution $f_{agh}(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$. On the other hand, when $s=0$ the approximation (13) is equal to one, that is the approximated expected value $E_\phi\left[\exp\{\nu(\mathbf{b})\}\right]$ is the unity, and $f_{dim}(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$ equals the Laplace approximation $f_L$. Hence, in this study, model estimation is obtained by maximizing the approximated log-likelihood obtained by replacing (15) into (7) for a given dimension $0\leq s<<d$. A quasi-Newton algorithm is used with gradient and Hessian matrix obtained numerically.

# 4 | ON THE DYNAMICS OF COGNITIVE ABILITIES AMONG ELDERLY IN THE UNITED STATES

We start our analysis by carrying out a preliminary investigation of the HRS/AHEAD data. Since the items are not continuous, Pearson correlations cannot be computed. Thus, we analyze the associations between pairs of items to highlight if it is worth fitting a dynamic factor model to the considered data. Indeed, significant associations between different items at the same and different occasions motivate the inclusion of the time-specific factors in the model as well as the analysis of their temporal relationships. On the other hand, significant associations between repeated measures of the same item justify the inclusion of the item-dependent random effects.[21] Table 4 reports the chi-square tests, for evaluating the significance of these associations, whose $p$-values are greater or equal to $10^{-4}$. The highest $p$-value is equal to .0044, indicating that all the items between and within time are indeed significantly related. Afterward, we have performed separate confirmatory factor analyses at each wave that have confirmed the bidimensionality of the seven cognitive items as found in previous studies.[5,29]

TABLE 4 Pairwise associations between items, HRS/AHEAD data

|  |  | Chi-square | df | *p*-value |
| --- | --- | --- | --- | --- |
| *PRES8* | *IR9* | 25.555 | 10 | .0044 |
| *YEAR8* | *PRES10* | 8.350 | 1 | .0039 |
| *PLANT8* | *SER7s9* | 18.931 | 5 | .0020 |
| *PLANT9* | *YEAR9* | 10.568 | 1 | .0012 |
| *PLANT8* | *SER7s9* | 20.473 | 5 | .0010 |
| *PLANT8* | *YEAR10* | 12.530 | 1 | .0004 |
| *SER7s8* | *PRES9* | 23.395 | 5 | .0003 |
| *PLANT8* | *PRES10* | 13.500 | 1 | .0002 |
| *PLANT8* | *VCPRES10* | 15.725 | 1 | .0001 |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

TABLE 5 BIC criterion for measurement invariance, $s = 3$, HRS/AHEAD data

| Models | loglik | ♯ par | BIC |
| --- | --- | --- | --- |
| Mod1 | −58 514.13 | 64 | 116 146.12 |
| Mod2 | −56 907.42 | 50 | 114 000.55 |
| Mod3 | −58 500.82 | 54 | 118 131.56 |
| Mod4 | −56 907.42 | 40 | 112 118.16 |
| Mod5 | −55 410.62 | 29 | 110 921.83 |
| Mod6 | −55 243.21 | 41 | 110 813.90 |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

## 4.1 | Results

First of all, we focus on the measurement part of the model and test if the assumption of measurement invariance of the intercepts and factor loadings (Assumption 2) holds, and if the random effects and time-specific latent factors fully account for the dependence among the observations (Assumption 1). At this stage, we do not make any assumption on the temporal relationship between the latent factors $\mathbf{z}_i$, and assume their covariance matrix to be unstructured.

We start by fitting the unconstrained and unstructured model defined by Equations (3) and (4), denoted as Mod1 in the following, characterized by item- and time-dependent thresholds $\tau_{jt}$ and loadings $\alpha_{jt}$, $j = 1, \ldots, 7, t = 8, 9, 10$. In all cases, the estimation is performed using the dimensionwise quadrature method with five quadrature points for dimension ($n_q = 5$) that results to be feasible by including up to four-dimensional integrals ($s = 4$). The choice of $s$ is done by first fitting the model with $s = 0$, that is just considering the leading term $f_L$ in Equation (9), and then increasing its value until the mean of the relative absolute differences in parameter estimates ($Av(\Delta_{s,s+1})$) is sufficiently small (order $10^{-3}$).[27] We have computed the values of this statistic for $s$ ranging from zero to four. There is a sensible change in the parameter estimates from $s = 0$ to $s = 1$ ($Av(\Delta_{01}) = 0.139$) and from $s = 2$ to $s = 3$ ($Av(\Delta_{23}) = 0.514$), whereas the estimates are quite stable from $s = 1$ to $s = 2$ ($Av(\Delta_{12}) = 0.056$) and from $s = 3$ to $s = 4$ that presents the smallest value ($Av(\Delta_{34}) = 0.006$). Hence, we consider $s = 3$ as the reference solution, and the dimensionwise quadrature based on the sum of uni-, bi-, and three-dimensional integrals is used to estimate all the subsequent models.

Measurement invariance over time is evaluated by fitting different models based on several constraints on the thresholds and factor loadings. Specifically, we consider the model characterized by time-invariant thresholds, that is $\tau_{jt} = \tau_j, t = 8, 9,$ and 10, denoted as Mod2, and also the one with equality constraints on the loadings over time, that is $\alpha_{jt} = \alpha_j$ for $t = 8, 9,$ and 10 (Mod3). Finally, we evaluate the model, called Mod4, in which both thresholds and loadings are assumed to be constant at each occasion. In order to assess the most appropriate invariance constraints for our data, all these models have been fitted and compared using the BIC criterion, reported in Table 5. The lowest BIC value is achieved for Mod4, indicating that the measurement invariance Assumption 2 is satisfied in these data.

**TABLE 6** Estimated model parameters with asymptotic standard errors in brackets, $s = 3$, HRS/AHEAD data

|  | $\hat{\alpha}_{j1}$ | $\hat{\alpha}_{j2}$ | $\hat{\sigma}^2_{u_j}$ |
|---|---|---|---|
| IR | 1.00 | – | 0.46 (0.26) |
| DR | 0.69 (0.16) | – | 0.73 (0.13) |
| SER7s | – | 1.00 | 1.16 (0.08) |
| PLANT | – | 0.92 (0.14) | 0.80 (0.47) |
| YEAR | – | 0.99 (0.11) | 0.81 (0.19) |
| PRES | – | 0.89 (0.14) | 0.77 (0.59) |
| VCPRES | – | 1.11 (0.12) | 0.83 (0.14) |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

**TABLE 7** Bivariate $\chi^2$-fits for pairs of items within time points, $s = 3$, HRS/AHEAD data

| Item | Wave 8 | | | | | | Wave 9 | | | | | | Wave 10 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IR | | | | | | | | | | | | | | | | | | |
| DR | 1.81 | | | | | | 2.05 | | | | | | 3.07 | | | | | |
| SER7s | 1.04 | 3.04 | | | | | 1.90 | 1.82 | | | | | 2.30 | 1.79 | | | | |
| PLANT | 2.31 | 1.81 | 3.81 | | | | 1.18 | 1.26 | 0.83 | | | | 2.14 | 1.79 | 1.83 | | | |
| YEAR | 3.18 | 3.63 | 3.56 | 4.05 | | | 1.36 | 3.37 | 1.99 | 2.79 | | | 3.41 | 2.01 | 2.59 | 1.95 | | |
| PRES | 2.50 | 2.22 | 4.81 | 3.87 | 2.29 | | 1.90 | 2.73 | 3.16 | 2.34 | 3.21 | | 3.19 | 2.35 | 2.73 | 0.61 | 1.19 | |
| VCPRES | 1.50 | 0.87 | 3.56 | 4.33 | 2.29 | 1.66 | 1.85 | 0.97 | 1.69 | 3.75 | 2.71 | 6.98 | 1.61 | 0.70 | 2.35 | 0.35 | 0.33 | 2.67 |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

Item-specific parameter estimates with asymptotic standard errors for Mod4 are reported in Table 6. The loading of the first item for each factor is fixed to one at each wave for the identification reasons previously discussed. The remaining estimates are all significant, confirming the two-factor structure of the data. Looking at the first latent variable, the episodic memory, we can notice that the loading associated to the delay recall *DR* item is smaller than the one associated to the immediate recall *IR*, indicating that the former discriminates less than the latter in this cognitive ability.

As for the second factor, the mental status, although all the loadings are very similar, we can observe that the most discriminating item is naming the Vice President (*VCPRES*). This result is in line with the descriptive analyses reported in Section 2.1 that have shown the worst performance of the spouses on this item with respect to all the other indicators of the individual mental status at each occasion. The estimated variances of the random effects associated to *IR* and *DR* are indicative of a small heterogeneity in their corresponding responses over time. On the other hand, the individuals' responses to the backward counting item *SER7s* are the most heterogeneous. The high fluctuation of the responses associated to this item was also highlighted for the period 1992 to 2004 in other studies.[5]

Finally, to evaluate if the assumption of conditional independence (Assumption 1) is satisfied in model Mod4, we have computed $\chi^2$-residuals on the bivariate contingency tables as goodness of fit measures rather than classical statistics that suffer from the sparseness problem.[43] The formula of the bivariate residuals, denoted with $\chi^2$-fits, is reported in the Appendix. Table 7 shows these statistics for all the two-way item combinations at each wave whereas the bivariate $\chi^2$-fits for pairs of items in different time points are reported in Table 8 of the Appendix. Each cell of the table is computed by summing the $\chi^2$-fits over all the response categories of the corresponding pair of items, and then by dividing this sum by the total number of possible category combinations. As a rule of thumb, values of the residuals greater than four are indicative of poor fit. Almost all the bivariate residuals are below this threshold, indicating that the time-dependent factors and random effects fully account for all the sources of variability between and within the items.

### 4.1.1 | Assessing the temporal relationships between the two cognitive abilities

In model Mod4, no assumptions are made on the temporal relationships between the time-dependent latent factors $\mathbf{z}_i$, whose covariance matrix is estimated to be unstructured as reported in Table 9. Analyzing the parameters related to the

**TABLE 8** Bivariate $\chi^2$-fits for pairs of items between time points, $s = 3$, HRS/AHEAD data

| Item | Wave 8,Wave 9 | | | | | | Wave 8, Wave 10 | | | | | | | Wave 9, Wave 10 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IR | 1.37 | 1.43 | 1.73 | 1.49 | 2.53 | 1.90 | 1.00 | 2.45 | 2.16 | 2.55 | 0.97 | 1.70 | 2.29 | 1.00 | 1.67 | 2.11 | 2.83 | 1.39 | 3.02 | 1.51 | 1.17 |
| DR | 0.90 | 2.40 | 2.87 | 1.33 | 2.88 | 1.91 | 1.76 | 1.59 | 2.45 | 2.40 | 1.42 | 3.55 | 2.53 | 1.76 | 1.29 | 1.38 | 1.36 | 2.11 | 3.54 | 2.67 | 0.68 |
| SER7s | 1.46 | 1.39 | 7.46 | 1.76 | 1.24 | 2.45 | 0.86 | 2.50 | 1.73 | 2.83 | 1.26 | 1.58 | 2.81 | 0.86 | 1.94 | 2.47 | 1.63 | 1.79 | 2.31 | 2.19 | 0.45 |
| PLANT | 2.56 | 3.29 | 5.85 | 0.92 | 3.17 | 1.87 | 3.03 | 3.39 | 3.09 | 2.94 | 1.58 | 4.00 | 3.53 | 3.03 | 2.20 | 2.75 | 1.42 | 1.04 | 3.09 | 2.85 | 0.54 |
| YEAR | 1.85 | 1.31 | 2.13 | 2.79 | 3.53 | 2.62 | 1.47 | 2.36 | 2.87 | 1.96 | 3.67 | 3.74 | 3.88 | 1.47 | 1.77 | 1.83 | 0.96 | 3.46 | 2.79 | 2.19 | 1.91 |
| PRES | 1.42 | 1.91 | 2.77 | 2.41 | 3.49 | 3.70 | 2.10 | 2.82 | 3.18 | 2.00 | 2.13 | 3.70 | 3.96 | 2.10 | 1.81 | 2.96 | 0.96 | 1.53 | 2.87 | 2.20 | 0.97 |
| VCPRES | 0.26 | 1.51 | 3.56 | 2.13 | 0.52 | 0.27 | 0.52 | 0.17 | 1.47 | 0.24 | 0.34 | 3.92 | 2.13 | 0.52 | 1.37 | 1.06 | 0.62 | 1.62 | 0.25 | 1.66 | 0.79 |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

$$
\Sigma_{\mathbf{z}} = \begin{pmatrix}
0.78(0.21) \\
0.14(0.07)\ 1.02(0.12) \\
0.76(0.20)\ 0.06(0.07) & 0.81(0.21) \\
0.17(0.07)\ 0.98(0.11) & 0.15(0.07)\ 0.99(0.12) \\
0.81(0.22)\ -0.01(0.07) & 0.81(0.22)\ 0.04(0.07) & 0.89(0.23) \\
0.26(0.07)\ 0.71(0.09) & 0.29(0.07)\ 0.72(0.09) & 0.26(0.07)\ 0.95(0.11)
\end{pmatrix}
$$

**TABLE 9** Estimated unstructured covariance matrix with standard errors in brackets, $s = 3$, HRS/AHEAD data

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

episodic memory ability, it can be noted that its variance increases over time, being equal to 0.78, 0.81, and 0.89 at wave 8, 9, and 10, respectively, indicating an increasing heterogeneity among individuals with respect to this dimension over time. This result can be due to differences among individuals who will develop AD vs those who will not, given the relative early decline in episodic memory in the preclinical AD individuals that makes task assessment of this cognitive domain particularly useful in the preclinical detection of at-risk individuals.[16] Furthermore, the covariances in the levels of this cognitive ability between adjacent waves result equal to 0.76 and 0.81, implying strong correlations (around 0.95) related to first-order autoregressive effects. Hence, as expected, the individual episodic memory ability at one occasion is strongly affected by the level achieved in the previous wave. This is also observed for the mental status ability, whose correlations between subsequent waves are equal to 0.98 and 0.72. However, differently from the episodic memory, the variance of this cognitive domain slightly decreases over time, being equal to 1.02, 0.99, and 0.95 at wave 8, 9, and 10, respectively.

Analyzing the temporal relationship between the two cognitive domains, we can observe that their covariance at each wave is quite low but always significant, being equal to 0.14, 0.15, and 0.26 at wave 8, 9, and 10, respectively. The corresponding correlations range from 0.16 to 0.28 and are indicative of weak synchronized effects between the two constructs. In terms of crosslagged relationships, we can observe that all the covariances between mental status at one occasion and episodic memory at the previous one are significant, and equal to 0.17 and 0.29. On the other hand, the covariances between episodic memory in one wave and mental status at the previous one are not significant, and equal to 0.06 and 0.04. In other words, episodic memory affects the mental status in subsequent occasions, but the converse does not hold. The predictive feature of episodic memory with respect to the more general cognitive domain is consistent with the studies that highlighted that deficits in episodic memory functioning occur very early in the AD disease process, whereas other cognitive domains are thought to decline later.[16,17]

The pattern of this unstructured covariance matrix appears to be consistent with a nonstationary first-order autoregressive process, being the variances of the two latent variables not constant over time and all the correlations related to the autoregressive effects strong and significant. Thus, we estimate a first-order autoregressive model as specified in Section 2.3.1, denoted as Mod5. It better fits the data with respect to the model Mod4, as indicated by the BIC criterion reported in Table 5.

The parameter estimates of this vector autoregressive model are provided in Table 10. The diagonal elements of the matrix $\mathbf{\Phi}_1$ represents the autoregressive coefficients of episodic memory (0.55) and mental status (1.04). Both

**TABLE 10** Structural parameter estimates with standard errors in brackets under Mod5, $s = 3$, HRS/AHEAD data

$$\boldsymbol{\Phi} = \begin{pmatrix} 0.55(0.03)\,0.05(0.13) \\ 0.36(0.14)\,1.04(0.65) \end{pmatrix} \quad \boldsymbol{\Sigma}_{\epsilon_1} = \begin{pmatrix} 0.33(0.05) \\ 0.39(0.04)\,0.47(0.06) \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{\epsilon} = \begin{pmatrix} 0.25(0.05) \\ -0.02(0.10)\,0.40(0.13) \end{pmatrix}$$

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

**TABLE 11** Estimated covariate regression coefficients, $s = 3$, HRS/AHEAD data

|  | Wave 8 | | Waves 9 and 10 | |
| --- | --- | --- | --- | --- |
|  | **Episodic memory** | **Mental status** | **Episodic memory** | **Mental status** |
| Age | −0.097 (0.008) | 0.012 (0.047) | −0.016 (0.006) | −0.155 (0.036) |
| Gender | −0.006 (0.007) | 0.037 (0.007) | −0.037 (0.006) | 0.046 (0.006) |
| Educ | −0.005 (0.047) | 0.089 (0.007) | −0.052 (0.040) | 0.039 (0.006) |

Abbreviation: HRS/AHEAD, Health and Retirement Study combined with the Asset and Health Dynamic study.

estimates are significant, but it is evident that mental status exhibits stronger persistence effects over time than episodic memory. This means that perturbations in the former cognitive domain continue to have a considerable impact in subsequent occasions. This is particularly relevant in view of previous studies on HRS cognitive measures that have highlighted the role of the mental status ability, whose indicators represent very simple mental processes capacity and alertness, in identifying serious individual deficiencies that standard cognitive tests cannot measure.[29]

Off-diagonal elements of $\boldsymbol{\Phi}_1$ represent crosslagged effects between the two cognitive abilities. In particular, the coefficient related to the influence of the episodic memory on the mental status at the subsequent occasion is equal to 0.36 and significant. On the other hand, there is no effect of mental status on episodic memory at the next occasion, being the corresponding crosslagged coefficient equal to 0.05, and not significantly different from zero.

Finally, the estimated covariance matrices $\boldsymbol{\Sigma}_{\epsilon_8}$ and $\boldsymbol{\Sigma}_{\epsilon}$ show a significant covariance just at the first wave, and not in the subsequent occasions, indicating that, after wave 8, there are not synchronized effects between the two cognitive abilities as was highlighted by Mod4.

## 4.1.2 | Effects of the covariates on cognitive functioning

We extend the vector autoregressive model Mod5 to also evaluate the effect of the three covariates, age (at 2006), gender and years of education, on both the latent variables in the three waves. This conditional model, denoted as Mod6, presents the best fit to our data, as shown in Table 5.

Table 11 shows the estimated regression coefficients at wave 8 and at the subsequent occasions (waves 9 and 10), when these parameters are assumed to be time-invariant. We can observe that age has a significant negative effect on the episodic memory at all the observed time points, whereas gender is significant only after the first occasion. This means that the episodic memory ability deteriorates with increasing age, and after the first wave it worsens for females with respect to males. On the other hand, age has a significant negative effect on the mental status ability only in waves 9 and 10, whereas gender and years of education positively affect this cognitive domain at all the occasions. This implies that a better performance on this ability is achieved by females and if the individual has a higher educational level.

The negative effect of age on the decline of both cognitive domains is consistent with previous studies on HRS/AHEAD data conducted on households' primary respondents[5,6,29] and, in general, with previous works on cognitive aging.[44] On the contrary, for the cohort of spouses considered in this study the effect of gender is opposite with respect to the previous analyses on the households' primary respondents, that is women performed better than men on mental status tasks but worse on memory tasks.

# 5 | DISCUSSION

In this article, we presented a latent variable model for multidimensional longitudinal data for the analysis of cognitive functioning in the older US population. The proposed model extends the existing literature by including multiple time-dependent latent variables, specifically episodic memory and mental status, and by assuming the nonstationarity of the cognitive domains over time. In particular, the dynamics of the two-dimensional cognitive process has been studied through a vector autoregressive model of first order that allows us to evaluate autoregressive, synchronized but also cross-lagged effects of its components. Possible extensions allow common factor scores from earlier occasions to also have direct effects. Higher order effects can be included in the model through the specification of the matrix $\boldsymbol{\Phi}_{t,t-j}, j = 2, \dots, l$, related to lagged vectors $\mathbf{z}_{it-j}$. The diagonal elements of $\boldsymbol{\Phi}_{t,t-j}$ are autoregressive effects, whereas higher order crosslagged effects are represented by its off-diagonal elements. The estimation of these models is unfeasible with traditional likelihood-based methods, because of the large number of latent variables and random effects needed to explain the different sources of variability present in the data. To overcome these computational problems, we considered the dimensionwise quadrature, that is based on a reduction of the dimensionality of the integrals involved in the likelihood function. It makes the estimation feasible in presence of a large number of latent variables/random effects, as in this study where we needed to approximate a 13-dimensional integral.

Our analysis on the HRS/AHEAD dataset focused on a sample of spouses of the households' primary respondents interviewed in the years 2006, 2008, and 2010. The two-factor model found in previous studies for the households' primary respondents has been confirmed in the three time points also for the spouses, with measurement invariance on the thresholds and loadings holding for both factors over time. Interestingly, we found that episodic memory predicts mental status over time. This result strengthens the importance of episodic memory in identifying and prevent cognitive impairment. The role of the socio-demographic variables age and education in predicting the cognitive performances has been widely investigated in the literature, highlighting the better performance of relatively young and well-educated individuals. In agreement with these previous findings, with reference to the cohort of spouses analyzed in this study, age is a clear predictor of both cognitive domains whereas level of education affects only mental status. Gender has a different effect on the two dimensions. Females have a better performance than males on mental status but are worse on episodic memory. Thus, it is important to take the demographic variables into account when cognitive tests are used for clinical purposes.

The proposed methodology has been discussed by considering balanced complete data. However, in some prospective studies, observations can be unequally spaced. Irregularly timed data can occur in two main situations: (a) the data are equally spaced with missing observations, and/or (b) the time distance between subsequent occasions is not necessarily constant, implying irregularly spaced assessment waves.

For unequally spaced observations, further extensions, especially for the vector autoregressive model, are needed. This can be tackled by providing a continuous time representation of the model. Continuous time processes can be specified using a broad class of differential equations, allowing for a wide degree of diversity in the types of dynamics that are being considered. A continuous time representation of the stationary VAR(1) process, known as Ornstein-Uhlenbeck (OU) process[45] has been recently proposed.[46] A similar specification has been derived for discrete-state hidden Markov processes.[47] A generalization that incorporates a latent linear mixed model with an OU process into the variance component has been also developed.[48] These extensions are important because discrete time processes can be sensitive to the choice of uniform time interval, that is different conclusions might be obtained if the autoregressive process is used for different time distances.[49,50] However, in these studies at most three latent variables have been considered at each occasion. Generalizations to more time-dependent latent constructs impact constraints placed on the drift matrix that specifies the autoregressive relationship in continuous time, and may not be easy to solve. The computations related to the matrix exponential operator applied to the drift matrix are demanding. Several approximations have been used in the literature, such as the Pade approximation,[51] oversampling techniques,[49] or the hybrid Kalman filter,[52] but still the computational time is long.

Alternative practical solutions can be used, without necessitating for the estimation of a continuous time model. For instance, the problem of unequally spaced measurements in discrete time modeling can be addressed by defining a time grid an adding missing data to the observations, to make the occasions approximately equally spaced in time. Some simulation studies have indicated that this largely reduces the bias that results from using discrete time estimation of unequally spaced data.[36,53]

Missing data represent a frequent issue in longitudinal studies, as participants may not be available at all the planned time occasions. The resulting data are unbalanced with an unequal number of measures for each subject. Monotone missingness (dropout) represents a frequent type of such nonparticipation, with some individuals leaving the study

prematurely and having a null probability to reenter. It contrasts with intermittent missingness where an individual who does not show up at a given occasion might return in a subsequent one. This latter is less common in cognitive functioning studies, where generally participants who leave the study do that permanently. In this context, a common assumption is also that, at each time point, variables for a respondent are either fully observed or totally missing; that is, there is no item nonresponse.

Let $\mathbf{W}_i = (W_{i1}, \ldots, W_{iT})'$ denote a $T$-dimensional vector of missing data, with $W_{it} = 0$ if the data $\mathbf{y}_{it}$ for the $i$th subject are available at time $t$, and equal to one otherwise. When modeling multidimensional longitudinal data subject to missing data, the joint density function of the complete data and latent variables given the covariates, $f(\mathbf{y}_i, \mathbf{W}_i, \mathbf{b}_i | \mathbf{x}_i)$, should be specified. It can be factorized as follows

$$f(\mathbf{y}_i, \mathbf{W}_i, \mathbf{b}_i | \mathbf{x}_i) = P(\mathbf{W}_i | \mathbf{y}_i, \mathbf{b}_i, \mathbf{x}_i) g(\mathbf{y}_i | \mathbf{b}_i, \mathbf{x}_i) h(\mathbf{b}_i | \mathbf{x}_i),$$

where $g(\mathbf{y}_i | \mathbf{b}_i, \mathbf{x}_i) h(\mathbf{b}_i | \mathbf{x}_i)$ refers to any model specification given in Section 2. As discussed in the article, the structural model $h(\mathbf{b}_i | \mathbf{x}_i)$ is the main focus of substantive research questions. $P(\mathbf{W}_i | \mathbf{y}_i, \mathbf{b}_i, \mathbf{x}_i)$ represents the probability associated to the missing data mechanism. A crucial distinction is whether the latter is missing at random or missing not at random (MNAR), that is, whether the probability of missingness does not or does depend on data which are themselves unobserved.[54,55] In the former case, if the parameters of the nonresponse model and those of the structural and measurement models are distinct from each other, the missingness in $\mathbf{y}_i$ is ignorable. This means that valid likelihood-based estimation and inference of the longitudinal model can be done by omitting the nonresponse model and performed as discussed in Section 3.[56] By contrast, MNAR is nonignorable, so the nonresponse model needs to be included in the analysis to obtain valid estimates and inference for the parameters of main interest. A common approach adopted in the latent variable literature is to introduce one or more additional latent variables $\boldsymbol{\xi}_i = (\xi_{i1}, \ldots, \xi_{iT})'$, called latent response propensities.[57-59] Generally $\xi_{it}, t = 1, \ldots, T$, is assumed to be time-invariant and categorical,[59] but we suggest to allow $\xi_{it}$ to be time-varying and continuous.[57] One advantage of the use of the latent response propensity model in the GLLVM framework is that it can be combined with any structural and measurement model described in Section 2, whereas the estimation of model parameters can be performed as described in Section 3.

The latent variable $\boldsymbol{\xi}_i$ represents unobservable determinants of the nonresponse, such that $\xi_{it}$ can be interpreted as the propensity of the $i$th individual to leave (or stay in) the study at the $t$th occasion. In this framework, all the associations between $(\mathbf{b}_i, \mathbf{x}_i)$ and $W_{it}$ are mediated entirely by $\boldsymbol{\xi}_i$, such that the hypothesis of ignorable missing data has the simple form that $\mathbf{b}_i$ and $\xi_{it}$ are conditionally independent. Different specifications of the relationships between the latent response propensity $\xi_{it}$ and the latent cognitive domains $\mathbf{z}_i$ have been proposed.[57] We suggest to allow the propensity of dropping out at a given occasion to depend on the value of the latent cognitive domains at the immediately preceding occasion. One advantage is that the hypothesis of ignorable nonresponse can be simply defined in terms of equality to zero of the structural parameters in the nonresponse model. An alternative formulation consists in allowing the response propensity $\xi_{it}$ to depend on $\mathbf{z}_1$ and covariates $\mathbf{x}_1$, both measured at the first time point.[57] The comparison among these different formulations is the topic of our current research.

A further potential extension of the proposed model, in particular when used for the analysis of cognitive decline, concerns its specification within a causal inference framework to also control for potential observed and unobserved confounders. This is particularly challenging since in presence of multiple (longitudinal) outcomes, there are several issues that have to be faced.

When outcomes of different nature are observed, different link functions have to be chosen and this can lead to different effects of the risk factors and controls for potential confounders on the various outcomes. In this case, the risk factors are included in the linear predictor, and their direct effects on the outcomes are analyzed. Furthermore, if the multiple longitudinal outcomes are measures of one or more time-dependent latent constructs at each occasion, they will show significant associations generated by the latent variables that have to be properly accounted for in a confirmatory dynamic factor model.

To perform correct causal inference in a multivariate longitudinal setting, and control for potential confounders, both these aspects must be considered. Several proposals have been discussed in the recent literature. A Bayesian dynamic factor model that allows to simultaneously estimate the causal effect of a treatment on the observed outcomes and account for the associations among them has been proposed.[60] The estimation of the average treatment effects is obtained by including in the model a first set of time-specific latent variables that also accounts for possible unobserved confounders. A second set of time-dependent latent variables is specified in the model to account for the associations between multiple outcomes. The temporal dependence between the latent variables is modeled by means of autoregressive processes. The

proposal is quite exhaustive and allows to solve several issues related to causal inference in multivariate longitudinal setting. However, all observed outcomes are assumed to be continuous.

Alternatively, a causal latent Markov model to evaluate the effects of one or more treatments on multiple longitudinal outcomes of different nature has been recently discussed.[61,62] As in the classical latent Markov model, the associations among the observed outcomes are explained by discrete time-dependent latent variables with two or more categories, and the temporal dependence among them is captured by estimating transition probabilities. Observed potential confounders are controlled by using a propensity score method. It replicates a pseudo-random assignment environment by producing weights that are included in the likelihood expression in a two-step estimation strategy. The novelty of this approach consists in evaluating the effect of one or more treatments directly on the latent variables, rather than in the measurement part of the model. This allows to avoid interpretation problems that can arise in presence of outcomes of different type.

Following this latter proposal, our model can be extended by first applying a propensity score method to control for potential observed confounders and then by evaluating the effects of risk factors on the two latent cognitive domains. Applying this approach could allow us to take into account for the association between observed variables by means of continuous latent variables and random effects as discussed in Section 2, and at the same time account for the effect of risk factors on the latent variables by controlling for potential confounders.

## DATA AVAILABILITY STATEMENT
The data are available at http://hrsonline.isr.umich.edu/

## ORCID
*Silvia Bianconcini* https://orcid.org/0000-0003-4810-5197
*Silvia Cagnone* https://orcid.org/0000-0002-7111-2945

## REFERENCES
1. McArdle J, Smith J, Willis RJ. Cognition and economic outcomes in the health and retirement study. In: Wise D, ed. *Explorations in the Economics of Aging*. Cambridge, MA: National Bureau of Economic Research; 2011.
2. He W, Goodkind DPK. *An Aging World: 2015*. U.S. Census Bureau, International Population Reports, P95/16-1. Washington, DC: U.S. Government Publishing Office; 2016.
3. Plassman B, Langa K, Fisher G, Heeringa S, Weir D, Ofstedal M. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*. 2007;29:125-132.
4. Plassman B, Langa K, Fisher G, Heeringa S, Weir D, Ofstedal M. Prevalence of cognitive impairment without dementia in the United States. *Ann Internal Med*. 2008;148:427-434.
5. McArdle J, Fisher G, Kadlec K. Latent variable analyses of age trends of cognition in the health and retirement study, 1992–2004. *Psychol Aging*. 2007;22:525-545.
6. Blankson A, McArdle J. A brief report on the factor structure of the cognitive measures in the HRS/AHEAD studies. *J Aging Res*. 2014;2014:1-5.
7. Li C, Hsu H. Cognitive function and associated factors among older people in Taiwan: Age and sex differences. *Arch Gerontol Geriatr*. 2015;60:196-200.
8. Ofstedal MB, Zimmer ZS, Lin HS. A comparison of correlates of cognitive functioning in older persons in Taiwan and the United States. *J Gerontol Soc Sci*. 1999;54:291-301.
9. Suthers K, Kim JK, Crimmins E. Life expectancy with cognitive impairment in the older population of the United States. *J Gerontol Soc Sci*. 2003;58:179-186.
10. Cagney KA, Lauderdale DS. Education, wealth, and cognitive function in later life. *J Gerontol Psychol Sci*. 2003;57:163-172.
11. Lievre A, Alley D, Crimmins EM. Educational differentials in life expectancy with cognitive impairment among the elderly in the United States. *J Aging Health*. 2008;20:456-477.
12. Anderlucci L, Viroli C. Covariance pattern mixture models for the analysis of multivariate heterogeneous longitudinal data. *Ann Appl Stat*. 2015;9(2):777-800.
13. Huang F, Zhang M, Wang S. Changes in cognitive function among older adults: a latent profile transition analysis. *Arch Gerontol Geriatr*. 2019;80:12-19.
14. Tromp DA, Lithfous T, Despres O. Episodic memory in normal aging and Alzheimer disease: Insights from imaging and behavioral studies. *Aging Res Rev*. 2015;24:232-262.
15. Small B, Bäckman L. Longitudinal trajectories of cognitive change inpreclinical Alzheimer's disease: a growth mixture modeling analysis. *Cortex*. 2007;43:826-834.
16. Bäckman L, Jones S, Berger A, Laukka E, Small B. Multiple cognitive deficits during the transition to Alzheimer's disease. *J Int Med*. 2004;256:195-204.

17. Small B, Herlitz A, Fratiglioni L, Almkvist O, Bäckman L. Cognitive predictors of incident Alzheimer's disease: a prospective longitudinal study. *Neuropsychology*. 1997;11:413-420.

18. Moustaki I, Knott M. Generalized latent trait models. *Psychometrika*. 2000;65:391-411.

19. Fieuws S, Verbeke G. Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach. *Stat Med*. 2004;23:3093-3104.

20. Dunson D. Dynamic latent trait models for multidimensional longitudinal data. *J Am Stat Assoc*. 2003;98:555-563.

21. Cagnone S, Moustaki I, Vasdekis V. Latent variable models for multivariate longitudinal ordinal responses. *Br J Math Stat Psychol*. 2009;62:401-415.

22. Pinheiro J, Bates D. Approximation to the loglikelihood function in the nonlinear mixed effects model. *J Comput Graph Stat*. 1995;4:12-35.

23. Rabe-Hesketh S, Skrondal A, Pickles A. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *J Econometr*. 2005;128:301-323.

24. De Bruijn N. *Asymptotic Methods in Analysis*. New York, NY: Dover Publications; 1981.

25. Huber P, Ronchetti E, Victoria-Feser M. Estimation of generalized linear latent variable models. *J Royal Stat Soc Ser B*. 2004;66:893-908.

26. Joe H. Accuracy of Laplace approximation for discrete response mixed models. *Comput Stat Data Anal*. 2008;52:5066-5074.

27. Bianconcini S, Cagnone S, Rizopoulos D. Approximate likelihood inference in generalized linear latent variable models based on dimension-wise quadrature. *Electron J Stat*. 2017;11:4404-4423.

28. Bollen K. *Structural Equations with Latent Variables*. New York: John Wiley and Sons; 1989.

29. Herzog AR, Wallace RB. Measures of cognitive functioning in the AHEAD study. *J Gerontol Ser B Psychol Sci Soc Sci*. 1997;52:37-48.

30. Jeon M, Rabe-Hesketh S. An autoregressive growth model for longitudinal item analysis. *Psychometrika*. 2016;81:830-850.

31. Andrich D. Controlling response dependence in the measurement of change using the Rasch model. *Stat Methods Med Res*. 2017;27:1-17.

32. Tran TD, Lesaffre E, Verbeke G, Duyck J. Modeling local dependence in latent vector autoregressive models. *Biostatistics*. 2021; 22(1):148-163.

33. Oort F. Three-mode models for multivariate longitudinal datl. *Br J Math Stat Psychol*. 2001;54:49-78.

34. Jöreskog K. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*. 1969;34:183-202.

35. Verbeke G, Fieuws S, Molemberghs G, Davidian M. The analysis of multivariate longitudinal data: a review. *Stat Methods Med Res*. 2014;23:42-59.

36. Cui K, Dunson DB. Generalized dynamic factor models for mixed-measurement time series. *J Comput Graph Stat*. 2014;23:169-191.

37. Duncan TE, Duncan SC, Strycker LA. *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications*. 2nd ed. New Jersey: Lawrence Erlbaum Associates Publishers; 2006.

38. McArdle J. Dynamic but structural modeling of repeated measures data. In: Nesselroade J, Cattell R, eds. *The Handbook of Multivariate Psychology*, New York: Springer; 1988.

39. Schilling S, Bock R. High-dimensional maximum marginal likelihood item factor analysis by adaptive quadrature. *Psychometrika*. 2005;70:533-555.

40. Bianconcini S. Asymptotic properties of adaptive maximum likelihood estimators in latent variable models. *Bernoulli*. 2014;20(3):1507-1531.

41. Magnus JR, Neudecker H. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Vol 393. Hoboken, NY: John Wiley & Sons.; 1988.

42. Liu Q, Pierce D. A note on Gauss-Hermite quadrature. *Biometrika*. 1994;81:624-629.

43. Jöreskog K, Moustaki I. Factor analysis of ordinal variables: a comparison of three approaches. *Multivar Behav Res*. 2001;36:347-387.

44. Hultsch D, Hertzog C, Dixon R, Small B. *Memory Change in the Aged*. New York, NY: Cambridge University Press; 1998.

45. Oravecz Z, Tuerlinckx F, Vandekerckhove J. Bayesian data analysis with the bivariate hierarchical Ornstein-Uhlenbeck process model. *Multivar Behav Res*. 2016;51:106-119.

46. Tran TD, Lesaffre E, Verbeke G, Duyck J. Latent Ornstein-Uhlenbeck models for Bayesian analysis of multivariate longitudinal categorical responses. *Biometrics*. 2020. http://dx.doi.org/10.1111/biom.13292.

47. Bartolucci F, Farcomeni A. A shared-parameter continuous-time hidden Markov and survival model for longitudinal data with informative dropout. *Statistics in Medicine*. 2019;38(6):1056-1073.

48. Tran T, Lesaffre E, Verbeke G, Molenberghs G. Serial correlation structures in latent linear mixed models for analysis of multivariate longitudinal ordinal responses. *Stat Med*. 2021;43:578-592.

49. Voelkle M, Oud J. Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes. *Br J Math Stat Psychol*. 2013;66:103-126.

50. Kuiper R, Ryan O. Drawing conclusions from cross-lagged relationships: re-considering the role of the time-interval. *Struct Equat Model Multidiscipl J*. 2018;25:809-823.

51. Arioli M, Codenotti B, Fassino C. The pade method for computing the matrix exponential. *Linear Algebra Appl*. 1996;240:111-130.

52. Driver C, Voelkle M. Hierarchical Bayesian continuous time dynamic modeling. *Psychol Methods*. 2018;23:774-799.

53. De Haan-Rietdijk M, Keijsers L, Hamaker E. Discrete- versus continuous-time modeling of unequally spaced ESM data. *Front Psychol*. 2017;8:18-49.

54. Rubin DB. Inference and missing data. *Biometrika*. 1976;63:581-592.

55. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Hoboken, NJ: Wiley; 2002.

56. Molenberghs G, Kenward MG. *Missing Data in Clinical Studies*. Hoboken, NJ: Wiley; 2007.

57. Hafez M, Moustaki I, Kuha J. Analysis of multivariate longitudinal data subject to nonrandom dropout. *Struct Equat Model Multidiscipl J*. 2015;22(2):193-201.

58. Kuha J, Katsikatsou M, Moustaki I. Latent variable modelling with non-ignorable item non-response propensity models for cross-national analysis. *J Royal Stat Soc Ser A*. 2018;181(4):1169-1192.

59. Marino M, Alfo M. Finite mixtures of hidden Markov models for longitudinal responses subject to drop out. *Multivar Behav Res*. 2020;55(5):647-663.

60. Samartsidis P, Seaman S, Montagna S, Charlett A, Hickman M, De Angelis D. A Bayesian multivariate factor analysis model for evaluating an intervention by using observational time series data on multiple outcomes. *J Royal Stat Soc Ser A*. 2020;183:1437-1459.

61. Bartolucci F, Pennoni F, Vittadini G. Causal latent Markov model for the comparison of multiple treatments in observational longitudinal studies. *J Educ Behav Stat*. 2016;41:146-179.

62. Tullio F, Bartolucci F. Evaluating time-varying treatment effects in latent Markov models: an application to the effect of remittances on poverty dynamics. Paper 91459; 2019.

## APPENDIX. BIVARIATE MARGINAL RESIDUALS

When data are sparse the classical goodness of fit statistics are not valid. In these cases, we can compute residuals from two-way marginal distributions that reveal if there are pairs of items responsible of the bad fit. Indicating with $nr$ the number of possible response patterns, the bivariate marginal residual for items $i$ and $j$ with categories $c_i$ and $c_j$, respectively is defined as Reference 43

$$\chi^2 - fit^{(ij)} = n \sum_a^{c_i} \sum_b^{c_j} \frac{(f_{ab}^{(ij)} - \hat{\pi}_{ab}^{(ij)})^2}{\hat{\pi}_{ab}^{(ij)}} \quad i = 1, \dots, p-1 \quad j = i+1, \dots, p$$

where $\hat{\pi}_{ab}^{(ij)} = \sum_{r=1}^{nr} y_{rs_i}^{(i)} y_{rs_j}^{(j)} \hat{\pi}_r$, $y_{rs_i}^{(i)} = 1$ and $y_{rs_j}^{(j)} = 1$ if $s_i$ is equal to category $a$ of item $i$ and $s_j$ is equal to category $b$ of item $j$, respectively, and 0 otherwise. $f_{ab}^{(ij)}$ are the correspondent observed frequencies. As a rule of thumb, a value of $\chi^2 - fit^{(ij)}/(c_i * c_j)$ greater than 4 indicates bad fit associated to the pair if items $i$ and $j$. In Table 8 the bivariate $\chi^2$-fits for pairs of items in different time points are reported.