Up in the 'longitudinal research' air: Lessons learned from real-world research and data analysis applications

Alejandro de la Torre-Luque
Department of Legal Medicine, Psychiatry and Pathology. Universidad Complutense de Madrid, Spain

1. State of the art

Longitudinal research has a long tradition in medical disciplines, such as Epidemiology, Psychology, Oncology or Gerontology. Researchers may benefit from longitudinal designs as developmental or maturational effects can be studied, as well as intra-individual variability sources can be controlled over time. However, longitudinal studies are not free from methodological and analytical challenges, such as dealing with elevated drop-out rates and intermittent data series; low predictive accuracy when distal factors are involved; or identifying person-specific profiles, underlying overall courses.

2. New perspectives and contributions

This symposium aims to provide some recommendations on longitudinal research design to deal with methodological challenges. Moreover, it intends to present some innovative analytical approaches for longitudinal data. The symposium comprises four interesting papers. First, Dr. Paula Fernandez's paper focuses on providing some recommendations to deal with data loss and drop-out phenomenon from the earliest stages of a longitudinal study: study design and task planning. Second, Mr. Benedikt Langenberg introduces an innovative analytical strategy to deal with complex nonlinear longitudinal trajectories: the multi-group latent growth components approach. The third paper is led by Mr. Pablo Fernandez-Cancer and is focused on dealing with processes that develop dually. Finally, Dr. Alejandro de la Torre-Luque presents an application of growth mixture modelling to study self-regulation developmental course and potential heterogeneous trajectories.

3. Research and practical implications

This symposium aims at filling some methodological and data analysis gaps in order to help researchers overcome challenging situations commonly seem in longitudinal studies, with a clear interest in real-world application.

Keywords: Longitudinal analysis; Missing data; Latent growth models; Latent class models.

E-mail: af.delatorre@ucm.es

What are we talking about when we talk about avoiding losing data? The spectrum of active intervention

M. P. Fernández,[1*] G. Vallejo,[1] E. Tuero-Herrero,[1] P. Livácic-Rojas,[2] and F. Ordóñez[3]

[1]University of Oviedo, Spain.

[2] University of Santiago, Chile.

[3] Faculty Padre Ossó of Oviedo and International University of La Rioja, Spain.

Abstract

Data loss is only harmless when it is completely random. And with this, we understand that the amount of lost data is insignificant and the responsible causes are improbable and unpredictable. If not, data loss has statistical and substantive consequences, to a greater or lesser extent depending on the amount of data lost, the underlying loss mechanism and the health status of the data that is complete. Assuming that the health status of the complete data is good, it is necessary to know how to use statistical data analysis techniques to cope with the loss of information, and then perform complex sensitivity analyzes to justify the inferences. But this is not the point at which we are going to focus this communication. We will return and be located in the anteroom of the investigation, we will stop actively in the planning. What we are going to do is prepare a plan to ensure the health of the investigation. We show the difference between two legitimate ways of planning an investigation, but only one of them will lead us to our objective.

Keywords: Data loss; Research planning; Active Intervention; Useful Rules; Substantive Sensitivity Analysis.

Funding: This research was supported by grant Ref.: PGC2018-101574-B-I00 (AEI/FEDER, UE) from the Spanish Ministry of Science, Innovation, and Universities.

***** *Corresponding author:* paula@uniovi.es

1. State of the art

In 1976 Rubin radiographed the footprint of the lost data. He gave them color, and described three types: MCAR, MAR and MNAR.

It is often said that data loss is only harmless when it is completely random, and with this we are assuming that the amount of data lost is negligible and the responsible causes are unlikely and unpredictable. However, this is not always the case, it depends on the particularities of the sample in which these losses occur (see Fernández, Vallejo, Livácic and Tuero, 2018). In any case, when the loss of data is otherwise, to a greater extent when the percentage of loss departs generously from that normative amount of 5%, the statistical and substantive consequences are inevitable. There are statistical options to deal with data loss, viz.

If the loss is MCAR or MAR: the information contained in the lost data is mostly contained in the complete data, and therefore it is possible to rescue it with statistical techniques such as LMM or IM.

If the loss is MNAR: there are statistical engineering techniques such as MPMM, SMSP, capable of taking advantage of the partial information that the complete data may contain. But it is not enough because the cause them, or is unknown, or the missing data depends on the values of the variable itself that are unknown.

It is well known that experts strongly affirm that there is no single or totally effective solution to deal with this problem. However, the effectiveness of these techniques depends not only on the loss mechanism and the amount of loss, but also on the health of the complete data, and the health of the complete data depends on the planning of the investigation and the care. assisted in the process of carrying it out. For this reason, experts warn us that the best solution is prevention.

2. Logistic solution. Prevention.

This communication sets out the difference between two correct ways of planning research, however, only one of them is adequate to guarantee the effectiveness of statistical techniques that address data loss.

In order to put ourselves in a real and habitual situation of investigation, we are going to focus on experimental or quasi-experimental longitudinal designs, which is why they are the most common designs in Health Sciences (Medicine, Psychology) and also in Education (in epidemiological studies, The most common are observational longitudinal designs), but also the ones that are most likely to cause data loss, if only because there is more time to allow for data loss. A longitudinal investigation studies a problem over time, and with the passage of time, independent variables and control variables produce losses, some intermittent and

others definitive, and due to this, in the sample there is a bias in selection, which is the greatest enemy of inference. If this is not anticipated during research planning, the situation may become critical. We have two design options, and both are good, but depending on what conditions. Let's see what they consist of.

2.1. First option. Conduct research in autopilot mode.

Research approach. Actions.

-Postulate research hypotheses related to the treatment of interest

-Select VDs and control variables

-Select the sample and calculate the sample size

Conduct of the the investigation. Actions

-Random assignment of subjects to groups (with or without other experimental control strategies)

-Guarantee the integrity of treatment

-Optimize the registration of VSs and control

Data analysis: Two possibilities:

a-No data loss

-the application of the appropriate statistical techniques in an optimal way will test if there is sufficient empirical evidence to test or reject the hypothesis

b-There is data loss

-If the loss is MCAR or MAR. An analysis of data with the proper techniques and a well-performed sensitivity analysis may avoid a disastrous forecast, but it will remain somewhat uncertain.

-If the loss is due to any of the circumstances highlighted under the MNAR data loss mechanism, we will be in the minefield immersed in a trap that is difficult to solve with a sensitivity analysis.

2.2. Second option. Carry out research in assisted mode.

Research approach. Actions

-Consider the possibility of losing data and the causes that can cause data loss

-Define what is lost data

-Establish a protocol that determines how to make the records and how to apply the treatment variable to continue the investigation in the subjects that require it

-Select the VDs (primary and secondary) and control variables, and other possible variables with different functions depending on the characteristics of the data that it is possible to lose.

-Select the sample and calculate the sample size with the forecast of data loss (see Vallejo, Ato, Fernández, Livacic-Rojas, & Tuero-Herrero, 2016; Vallejo, Ato, Fernández, & Livácic, 2018).

-Postulate research hypotheses related to the treatment of interest.

-Postulate other possible research hypotheses related to data loss

Conduct of the investigation. Actions

- Implement Solutions to avoid losing data at the same time that we test the hypotheses defined above.

- Amplify observation opportunities:

-extend the registry adapted to the different solutions taken in order to evaluate different variations of the initial independent variable.

-Implement a monitoring protocol that allows evaluating different aspects that can only be observed over time.

Data analysis: Two possibilities:

a-No data loss

-the application of the appropriate statistical techniques in an optimal way will test if there is sufficient empirical evidence to test or reject the hypothesis

b-There is data loss

-If the loss is MCAR or MAR. A data analysis with the proper techniques and a sensitivity analysis will avoid the ruin of experimental research, and the consequences and the prognosis will be known.

-If the loss is due to any of the circumstances highlighted under the MNAR data loss mechanism, the experimental research will not be ruined, because we will have the possibility of converting the MNAR loss into MAR loss

3. Conclusions. The moral of the story

-When a research with a causal purpose that is carried out in the health sciences (eg Medicine, Psychology) or in other sciences, is carried out in automatic pilot mode, a "scientific anesthesia" is produced, incapable of responding to the loss of data if These are produced obeying an MNAR loss mechanism. It places the analysis of the data in an EMERGENCY situation, and therefore a difficult and uncertain solution, preventing the progress of the investigation.

-When an investigation with a causal purpose that is carried out in the health sciences (eg Medicine, Psychology) or in other sciences, is carried out in assisted mode, the investigation is conducted while being "ON GUARD", enabling possibilities that prevent the loss of information that is decisive in coping with the loss of data that occurs due to MNAR mechanisms. Under these design conditions, coupled with the help of more effective statistical techniques for dealing with missing data, it is possible that sensitivity analysis provides useful truths rather than establishing obscure and doubtful relationships. That is, we will know more accurately what the effect of the treatment has been, the lack of effect, etc.

In conclusion, we will know what causes selection bias, we will reduce the uncertainty caused by data loss, and we will increase the degree of confidence in the conclusions.

References

Fernández, P., Vallejo, G., Livácic-Rojas, P., & Tuero-Herrero, E. (2018). The (ir)responsibility of (under)estimating missing data. *Frontiers in Psychology, 9,* Article 556. https://doi.org/10.3389/fpsyg.2018.00556

Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63, 581–592. https://doi.org/10.1093/biomet/63.3.581

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

Vallejo, G., Ato, M., Fernández, P., and Livácic-Rojas, P. (2018). Sample size estimation for heterogeneous growth curve models with attrition. *Behavior Research Methods, 51,* 1216-1243. https://doi.org/10.3758/s13428-018-1059-y

Differences in Longitudinal Trajectories between Groups - The Multi-Group Latent Growth Components Approach

Benedikt Langenberg,* Axel Mayer
RWTH Aachen University, Germany

Abstract

Purpose: In this article, we propose a multi-group approach for analyzing complex nonlinear longitudinal trajectories. ***Method***: The approach is based on the latent growth components approach (LGCA) that offers a flexible framework for defining growth components and extends the same for the use with multiple groups. The approach benefits from known advantages of the LGCA and adds more capabilities from the multi-group framework, that is, (1) it can flexibly include complex nonlinear growth components, (2) incorporate a measurement model for the latent state variables and latent covariates, (3) it can model differences in growth components based on categorical covariates, and (4) treat covariates and group weights as fixed or stochastic. ***Results and conclusions***: We demonstrate the approach using data from the Health and Retirement Study that includes individuals diagnosed with cancer. We analyze trajectories in depressive symptoms before and after the cancer diagnosis with respect to a subset of categorical covariates (i.e., groups). We further present the open-source R package semnova that implements the proposed approach and makes it conveniently accessible for applied researchers.

Keywords: Latent growth models; Longitudinal research; Average effects; Multi-group analysis; Latent growth components approach.

Funding: This work was supported by the German Research Foundation under Grant MA 7702/1-1. The data that is used in this article was funded by the American National Institute on Aging under Grant NIA U01AG009740.

Supplemental material: The software code that supports the findings of this study is openly available in github at https://github.com/langenberg/LangenbergMayer2020_EAM2020.

***** *Corresponding author:* langenberg@psych.rwth-aachen.de

1. Introduction

A strong demand exists in research for the analysis of complex trajectories of change over time. For instance, the study of change in patients' well-being measured mupltiple times before and after a cancer diagnosis is of great interest in medical research. A number of methods has been proposed to address this challenge including *latent growth curve models* (McArdle, 1988; McArdle & Epstein, 1987; Meredith, 1993) and *latent change score models* (McArdle, 2009; McArdle & Hamagami, 2001; Raykov, 1999; Steyer, Eid, & Schwenkmezger, 1997). Latent growth curve models oftentimes aim at modeling polynomial trajectories of change and latent change score models focus on modeling the change between two neighboring measurement occasions. Reseachers, however, may have very particular hypotheses about the shape of change in patients' well-being. The *latent growth components model* (LGCA, Mayer, Steyer, & Mueller, 2012) is a generaliziation of the aforementioned models and staisfys this need offering the researcher a convient way to model complex trajectories of change.

All of the aforementioned models have in common that they were originally formulated as single-group models. That is, groups and categorical covariates are included in the model as dummy-coded predictors (e.g., Mayer, Geiser, Infurna, & Fiege, 2013). In this article, we extend the LGCA for the use with multiple groups to model effects of categorical covariates (*multi-group latent growth components approach*, MG-LGCA). For this purpose, we build on causality theory to estimate the average of the effects of categorical covariates which is conceptually similar to the *EffectLiteR approach* (Mayer, Dietzfelbinger, Rosseel, & Steyer, 2016). We demonstrate the MG-LGCA by means of data from the Health and Retirement Study containing physical and emotional depressive symptoms of patients before and after a cancer diagnosis as well as multiple categorical and continuous predictors. We employ the same model and the same data that was used by Mayer et al. (2013), but use a multi-group model instead of dummy-coded categorical covariates. We conclude this article by briefly discussing findings from the analysis.

2. Method

2.1. Motivating Example

For this article, we use the same data from the Health and Retirement Study ($N = 2{,}798$, University of Michigan, 2020) that was used by Mayer et al. (2013). The data contains a number of items measuring depressive symptoms thate were divided into two parcels phsyical symptoms (included items: *Felt depressed, Effort*, *Sleep*, *Not get going*) and emotional symptoms (included items: *Happy, Lonely, Enjoy life, Sad*). For each of the variables, two measurements before the cancer diagnosis, two measurements after the diagnosis as well as one measurement in the year of the diagnosis were selected (more

occasions are available, see, e.g., Infurna, Gerstorf, & Ram, 2013). The data consequently includes five measurement occasions that are two years apart from each other. For illustration, the patients' gender (G, female vs. male) and marital status (M, married or partnered vs. not married or partnered) were used as categorical predictors.

2.2. The Single-Group Latent Growth Components Model

We use the multi-state model from Mayer et al. (2013) to build the MG-LGCA upon. The LGCA enables the researcher to define custom contrasts (i.e., growth components) by specifying a contrast matrix $\mathbf{C}$ that transforms the latent state variables $\boldsymbol{\eta}$ into $\boldsymbol{\pi}$:

$$\pi = \mathbf{C}\eta \ .$$

This transformation cannot directly be implemented in SEM because the relevant part of the structural model takes the form $\boldsymbol{\eta} = \mathbf{B}^*\boldsymbol{\pi}$. To obtain the $\mathbf{B}^*$ matrix, the $\mathbf{C}$ matrix must be inverted:

$$\eta = \mathbf{C}^{-1}\pi = \mathbf{B}^*\pi \ .$$

The $\mathbf{B}^*$ matrix is then incorporated into the matrix of structural coefficients of the SEM containing regressions between the latent variables (see, Mayer et al., 2012, for details). Mayer et al. (2013) used this approach to formulate five growth components of change in patients' well-being before and after a cancer diagnosis. $\boldsymbol{\eta}$ corresponds to the latent state variables at the five measurement occasions measured by two items each (i.e., the two parcels). $\boldsymbol{\pi}$, on the other hand, corresponds to the growth components. The five components of interest were: (1) initial level π_0 at the first measurement occasion; (2) linear change component π_1; (3) reaction component π_2 defined as the difference of between the average of the two measurement occasions before the diagnosis and the measurement in the year of the diagnosis; (4) adaptation component π_3 defined as the difference between the two measurement occasions before the diagnosis and the two measurement occasions after the diagnosis; (5) post-diagnosis level π_4 defined as depressive symptoms at the first occasion after the diagnosis. The contrast matrix $\mathbf{C}$ and the corresponding inverse $\mathbf{B}^*$ are given by:

$$\mathbf{C} = \begin{array}{c} \\ \text{growth components} \end{array} \begin{array}{c} \text{latent state variables} \\ \begin{array}{ccccc} \eta_1 & \eta_2 & \eta_3 & \eta_3 & \eta_3 \end{array} \end{array} \begin{array}{c} \\ \begin{array}{c} \pi_0 \\ \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{array} \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 & 2 \\ -1 & -1 & 2 & 0 & 0 \\ -1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{B}^* = \mathbf{C}^{-1} = \begin{array}{c} \\ \text{latent state variables} \end{array} \begin{array}{c} \text{growth components} \\ \begin{array}{ccccc} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \pi_4 \end{array} \end{array} \begin{array}{c} \\ \begin{array}{c} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \end{array} \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0.5 & 0.5 & 0.5 & -1 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

Mayer et al. (2013) used a τ-equivalent measurement model for the two parcels at each measurement occasion fixing the intercepts of the parcels to zero and the loadings to one. An additional method factor was used to account for method effects of the parcels. For identification purposes, intercepts and residual (co-)variances of the $\boldsymbol{\eta}$ variables were fixed

to zero while means and covariances of $\boldsymbol{\pi}$ were freely estimated. The model fit reported by Mayer et al. (2013) was $\chi^2(28) = 87.979, p < .001$, $CFI = 0.989$, $RMSEA = 0.027$, $SRMR = 0.021$.

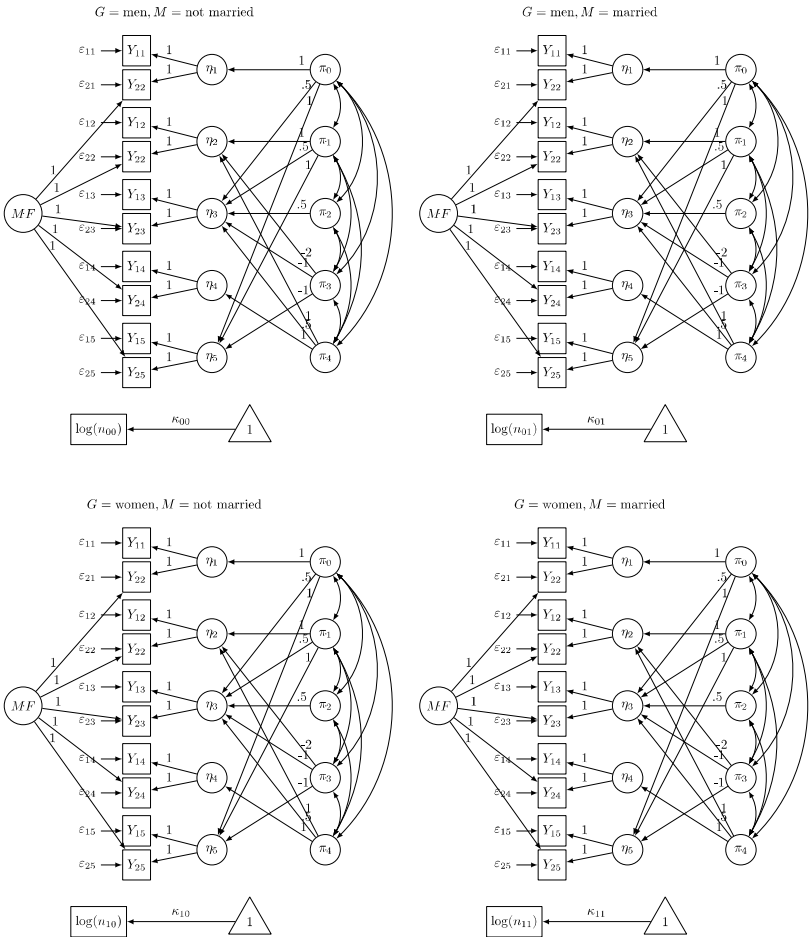

Figure 1. Complete model with latent state variables (η_t)*, growth components* (π_j) *and method factor* (MF) *separated by the four permutations of the categorical predictors* (G, M)*. For the sake of readability, the covariances between the growth components* (π_j) *and the method factor have been omitted.*

2.3. The Multi-Group Latent Growth Components Approach

In the MG-LGCM, separate models for every possible permutation of the categorical covariates are simultaneously estimated. The measurement model is invariant across the groups as well as the structural coefficients regressing the dependent variables $\boldsymbol{\eta}$ onto the growth components $\boldsymbol{\pi}$ and the means of the method factor. In this example, only an intercept is estimated for every growth component $\boldsymbol{\pi}$:

$$E(\pi_j|G = \mathsf{g}, M = \mathsf{m}) = \alpha_{j\mathsf{gm}} \ .$$

$\alpha_{j\mathsf{gm}}$ represents the mean of π_j in the corresponding group. Based on the parameter estimates of the separate models, a regression can be formulated for each of the $\boldsymbol{\pi}$ variables onto the categorical covariates:

$$E(\pi_j|G, M) = \beta_{j00} + \beta_{j1\_} \cdot I_{G=1} + \beta_{j\_1} \cdot I_{M=1} + \beta_{j11} \cdot I_{G=1} \cdot I_{M=1}$$

$I_{G=1}$ is an indicator variable that equals one of gender equals women. $I_{M=1}$ ist an indicator variable that equals one of marital status equals married. This regression is very similar to a single-group regression with dummy-coded categorical covariates and the coefficients are to be interpreted in the same way. β_{j0} is the mean of π_j in the group of men that are not married which serves as reference group. $\beta_{j1\_}$ represents the increase in π_j for not married women compared to not married men. β_{j_1} is the increase in π_j for married men compared to not married men. β_{j11} represents the interaction of gender and marital status. We further define the average effects of the categorical covariates using gender as an example by:

$$\mathsf{AVE}_{G;\pi_j} = E[E(\pi_j|G = \mathsf{women}, M) - E(\pi_j|G = \mathsf{men}, M)]$$

The average effects of the categorical covariates are defined as the average of the group-specific effects weighted by the group probabilities. For the average effect of gender, these group probabiltities correspond to the unconditional distribution (i.e., unconditional probability) of the categorical covariate marital status (M). The group probabilities are es well estimated from the data and treated as stochastic which is very similar to the EffectLiteR approach (Mayer, Dietzfelbinger, Rosseel, & Steyer, 2016).

3. Results

With the regression for each of the growth components $\boldsymbol{\pi}$, it is now possible to calculate the (conditional) expectations and average (conditional) effects of interest. For this demonstration, we focus on the average effects of the categorical covariates and the unconditional expectation of the π_j variables. We estimate the model using the SEM R software package lavaan (Rosseel, 2012) with full maximum likehood to account for missing values and a robust estimator. Table 1 contains the estimated regression coefficients $\boldsymbol{\beta}$, the average effects of gender and marital status and the unconditional expectation for each of the $\boldsymbol{\pi}$ variables of the MG-LGCA. The model fit is $\chi^2(115) = 232.277, p < .001$, $CFI = 0.981$, $RMSEA = 0.042$, $SRMR = 0.036$. Although, the χ^2-statistic is significant, the other fit indices are fairly good. Compared to the single-group multi-state model, the fit has slightly gotten worse, but can still be considered comparable. Figure 2 shows the model implied means of the five measurement occasions for each combination of the categorical covariates. From Table 1 and Figure 2, it can be seen that married man have the lowest baseline of

depressive symptoms π_0 ($\hat{\beta}_{000} + \hat{\beta}_{0\_1}$). There is, furthermore, a significant average effect of gender ($\widehat{\text{AVE}}_{G;\pi_0}$) and martial status ($\widehat{\text{AVE}}_{M;\pi_0}$) indicating that women show a higher baseline as well as not married participants. Married participants have on average a steeper linear trend π_1 ($\widehat{\text{AVE}}_{M;\pi_1}$) compared to not married participants. For the reaction growth component π_2, there is a significant average effect for marital status ($\widehat{\text{AVE}}_{M;\pi_2}$). Married participants show a greater reaction component. The interaction between gender and marital status ($\hat{\beta}_{211}$) is as well significant for the reaction component. For the adaptation component π_3, there is again a significant average effect for martial status ($\widehat{\text{AVE}}_{M;\pi_3}$). Married participants have higher depressive symptoms after the diagnosis while symptoms are the highest for married men ($\hat{\beta}_{300} + \hat{\beta}_{3\_1}$). The post diagnosis growth component π_4 is the smallest for married men ($\hat{\beta}_{400} + \hat{\beta}_{4\_1}$) which is, however, no surprise as this group has the lowest overall depressive symptoms. The post diagnosis level is on average lower for married participants ($\widehat{\text{AVE}}_{M;\pi_4}$) and lower for men ($\widehat{\text{AVE}}_{G;\pi_4}$).

Table 1. Regressions with predictors for the multi-group latent growth components model.

		π_0		π_1		π_2	
		Estimate	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
$\hat{E}(\pi_j)$		2.25*	0.07	1.50*	0.21	1.07*	0.13
$\widehat{\text{AVE}}_{G;\pi_j}$		0.30*	0.10	-0.33	0.32	0.16	0.21
$\widehat{\text{AVE}}_{M;\pi_j}$		-1.11*	0.12	1.07*	0.39	0.69*	0.23
$\hat{\beta}_{j00}$	(intercept)	2.92*	0.17	0.77	0.59	0.71*	0.33
$\hat{\beta}_{j1\_}$	(women)	0.16	0.21	0.12	0.69	-0.47	0.39
$\hat{\beta}_{j\_1}$	(married)	-1.20*	0.18	1.38*	0.62	0.26	0.35
$\hat{\beta}_{j11}$	(women, married)	0.20	0.23	-0.67	0.77	0.93*	0.46

		π_3		π_4	
		Estimate	*SE*	*Estimate*	*SE*
$\hat{E}(\pi_j)$		0.92*	0.12	2.69*	0.08
$\widehat{\text{AVE}}_{G;\pi_j}$		-0.23	0.20	0.27*	0.11
$\widehat{\text{AVE}}_{M;\pi_j}$		0.74*	0.24	-0.70*	0.13
$\hat{\beta}_{j00}$	(intercept)	0.40	0.35	3.02*	0.19
$\hat{\beta}_{j1\_}$	(women)	0.15	0.41	0.36	0.23
$\hat{\beta}_{j\_1}$	(married)	1.00*	0.37	-0.64*	0.20
$\hat{\beta}_{j11}$	(women, married)	-0.56	0.46	-0.14	0.26

Note: *$p < .05$.

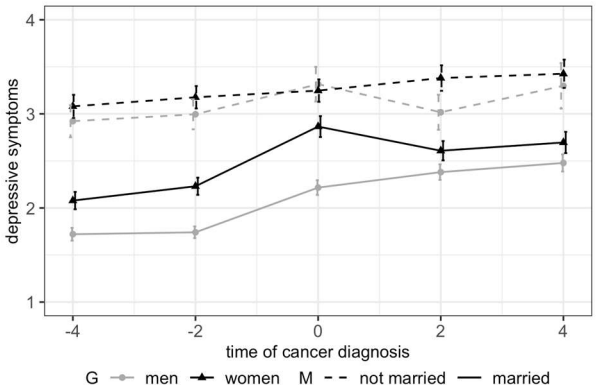

Figure 2. Model implied mean of drepressive symptoms for the for categorical covariates. Time zero corresponds to the year of the diagnosis. Negative values correspond to years before the diagnosis and positive values after the diagnosis, resprectively. Error bars indicate standard errors.

4. Conclusions

In this paper, we presented the multi-group extension to the latent growth components model. Using data from the Health and Retirement Study, we showed how to specify the MG-LGCA model for five growth components and two categorical covariates with two levels each.

The open-source R packge semnova (https://github.com/langenberg/semnova) implements the LGCA and makes the analysis of complex custom growth components conviently accessible to applied researcher. The semnova package was originally developed for latent repeated measures analysis of variance which is closely related to latent growth curve modeling and the latent growth components approach, and builds upon the LGCA for estimating the model. The package currently supports only the single-group LGCA, but the multi-group extension will soon be available.

References

Infurna, F. J., Gerstorf, D., & Ram, N. (2013). The nature and correlates of change in depressive symptoms with cancer diagnosis: reaction and adaptation. *Psychology and aging, 28*(2), 386–401. https://doi.org/10.1037/a0029775

Mayer, A., Dietzfelbinger, L., Rosseel, Y., & Steyer, R. (2016). The EffectLiteR Approach for Analyzing Average and Conditional Effects. *Multivariate Behavioral Research, 51*(2–3), 374–391. https://doi.org/10.1080/00273171.2016.1151334

Mayer, A., Geiser, C., Infurna, F. J., & Fiege, C. (2013). Modelling and predicting complex patterns of change using growth component models: An application to depression trajectories in cancer patients. *European Journal of Developmental Psychology, 10*(1), 40–59. https://doi.org/10.1080/17405629.2012.732721

Mayer, A., Steyer, R., & Mueller, H. (2012). A General Approach to Defining Latent Growth Components. *Structural Equation Modeling, 19*(4), 513–533. https://doi.org/10.1080/10705511.2012.713242

Mayer, A., & Thoemmes, F. (2019). Analysis of Variance Models with Stochastic Group Weights. *Multivariate Behavioral Research, 54*(4), 542–554. https://doi.org/10.1080/00273171.2018.1548960

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561–614). New York, NY: Plenum Press.

McArdle, J. J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. Annual Review of Psychology*, 60*(1), 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McArdle, J. J., & Epstein, D. B. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58*(1), 110–133.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. *New Methods for the Analysis of Change*. https://doi.org/10.1037/10409-005

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1 - 36. doi:http://dx.doi.org/10.18637/jss.v048.i02

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling True Intraindividual Change: True Change as a Latent Variable. *Methods of Psychological Research, 2*(1), 21–33.

Raykov, T. (1999). Are simple change scores obsolete? An approach to studying correlates and predictors of change. *Applied Psychological Measurement, 23*(2), 120–126. https://doi.org/10.1177/01466219922031248

University of Michigan. (2020). Health and Retirement Study [Data file]. Retrieved from https://hrs.isr.umich.edu

Effectiveness of the Bivariate Dual Latent Change Score model for longitudinal research

Pablo Fernández-Cáncer,* Eduardo Estrada
Faculty of Psychology, Universidad Autónoma de Madrid, Spain.

Abstract

Purpose. The Bivariate Dual Change Score (BDCS) model (McArdle, 2001) is a SEM useful for the study of two variables that unfold over time. Due to its flexibility to characterize change, stability, and lead-lag relations between latent processes over time, it is very frequent in developmental studies. Despite its popularity, very few studies have examined the effectiveness of this model to retrieve information from dynamic processes. To the best of our knowledge, existing literature focused on specific situations such as the impact of incomplete data (McArdle & Hamagami, 2001), or misspecification (Ji & Chow, 2018). Our general purpose is to evaluate the ability of the BLCS model to recover the characteristics of a dynamic longitudinal bivariate process under a broad range of empirically relevant conditions. *Method and design*. Through a Monte Carlo simulation, we manipulated the populational parameters and created various datasets defining different developmental trajectories. We also manipulated the sample size and the number of measurement occasions. We fitted BDCS models to the data and assessed the proportion of improper solutions and the recovery of the parameters under the different conditions. *Results*. With three measurement occasions, the BDCS model consistently led to higher bias and variability in the model parameters, as well as higher rates of improper solutions. With seven or more repeated measures, results were excellent regardless of the sample size and generating process. *Conclusions*. Based on our findings, we provide specific recommendations for the design of longitudinal studies, and the estimation of BDCS models in such studies.

Keywords: structural equation models; latent change score model; longitudinal data analysis; dynamical systems.

***** *Corresponding author:* pablofc_1995@hotmail.com

1. Introduction

Most developmental theories are multivariate and dynamic, inasmuch they describe processes that unfold over time and involve multiple interrelated parts. Consider, for example, the associations between development of cognitive functions (e.g. inhibition or working memory) and maturation of the cortical structure during childhood and adolescence. Does cortical maturation precede changes in cognitive development? Can cortical maturation be partly driven by cognitive development? How are these processes interrelated over time? The common goal of these questions is to identify and examine the sequences of time-lagged change in multiple variables. To evaluate this type of hypotheses, the methods used must be able to capture the patterns of interplay between processes over time. Latent Change Score (LCS) models (McArdle, 2001, 2009; McArdle & Hamagami, 2001) are a useful and popular approach to study dynamics in longitudinal data.

1.1. Latent Change Score Model (LCS)

LCS models incorporate a measurement structure that defines observed variables as a function of a true unobserved score plus measurement error. This makes it possible to model change as a latent variable representing the quantitative differences between true scores at adjacent time points. Within this framework, bivariate versions of these models (BLCS) have been developed to examine the interrelations between two variables that unfold over time (McArdle, 2001), allowing for a better understanding of multivariate processes. In a common specification of BLCS models, latent changes in two processes x and y (Δx and Δy), at any given time *t*, are function of three components (Ferrer & McArdle, 2010): (a) an additive component α, typically representing a constant influence on the system; (b) a self-feedback parameter β, representing the influence of the same variable at the previous occasion, *t*-1; and (c) a coupling parameter γ, representing the influence of the other variable at *t*-1. Additionally, LCS models allow the inclusion of prediction error at the latent level. These prediction errors are often termed *dynamic errors* (d_x and d_y), and they account for unobserved events such as fatigue, anxiety, or motivation, that are carried over through the time-lagged effects to latter states of the system (for more details, see Ji & Chow, 2019; Schuurman et al., 2015). Therefore, the equations for change at time *t* are expressed as:

$$\Delta x_{[t]} = \alpha_x \times x_a + \beta_x \times x_{[t-1]} + \gamma_x \times y_{[t-1]} + d_x$$

$$\Delta y_{[t]} = \alpha_y \times y_a + \beta_y \times y_{[t-1]} + \gamma_y \times x_{[t-1]} + d_y$$

where x_a and y_a are the additive components that influence the system trough α (which are often fixed to 1). Figure 1 depicts a path diagram for a BLCS.

The previous equations specify a linear dynamics of change. However, at every time point, self-feedback and coupling parameters are multiplied by scores at the previous state. This

leads to compound effects that can result in distinct nonlinear trajectories. This accumulation of effects allows lagged relations to capture both the short-term dynamics and the long-term developmental trajectories (Usami et al., 2019). Furthermore, multiple specifications are possible depending on the model of change that characterizes the dynamic process under study (e.g. dynamic error or *innovations* at the latent change level, time-variant lagged effects, or additional exogenous predictors). Nevertheless, the main interest of the BLCS model is the study of the interrelation between two processes over time. The already mentioned coupling parameters capture these interrelations and allow us to test hypotheses about the presence and sign of the interrelation, as well as the identification of the leading process (i.e. predictor of subsequent changes).

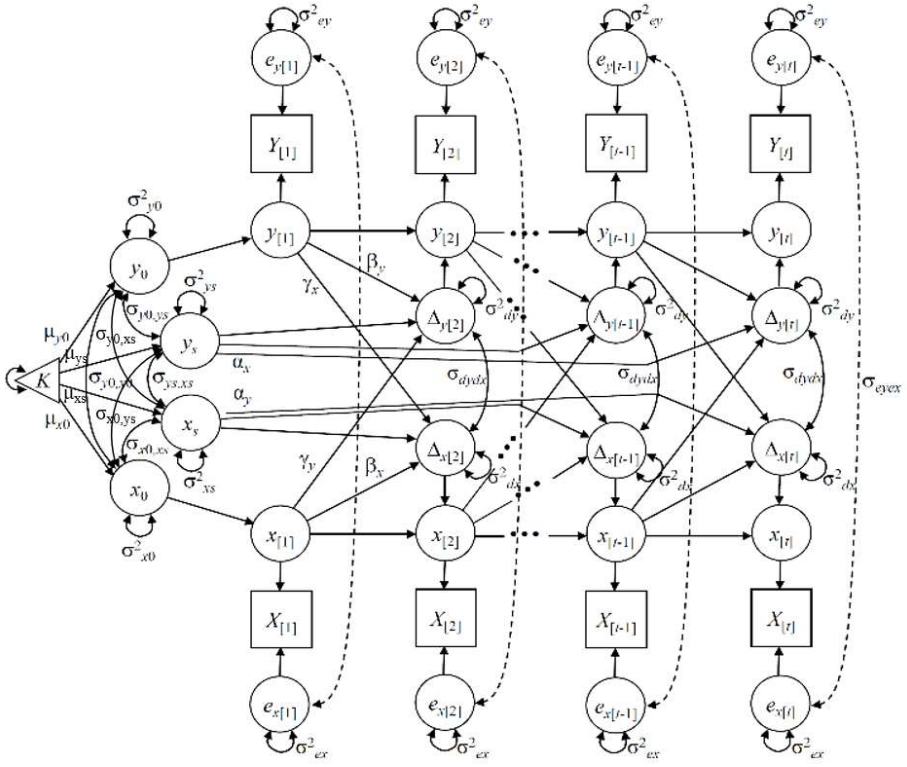

Figure 1. Path diagram of a BLCS

Due to their flexibility to characterize change, stability, and lead-lag relations between bivariate latent processes, BLCS models have become frequent in many areas of behavioural and health sciences. For example, studies on cognitive development during adulthood and old age indicate that processing speed is a leading indicator of age changes in memory and spatial ability, but not verbal ability (Finkel et al., 2007), and is predictive of later changes in knowledge but not vice versa (Ghisletta & Lindenberger, 2003). BLCS models make it

possible to examine complex dynamical hypothesis. For example, a recent study by Estrada et al., (2019) found that cortical and cognitive changes were related to each other reciprocally and the rate of change (not the actual level) was predictive of later changes. BLCS models are also used frequently in the study of psychological processes in dyads. For example, they have been used to study depressive symptoms in spouses (Kouros & Cummings, 2010), the relation between marital satisfaction and self-rated health (Proulx & Snyder-Rivas, 2013), and the effect of self-directed interventions on couples' communication (Bodenmann et al., 2014), among others (for more examples, see Estrada et al. 2020). BLCS have been applied also in psychopathology and clinical psychology (Kertes et al., 2008; King et al., 2006).

Despite the popularity and broad application of the BLCS model, very few studies have examined their effectiveness to retrieve information from dynamic processes. Namely, the ability of this model to recover the true populational parameters under different scenarios has not been sufficiently investigated. To the best of our knowledge, existing literature has focused on specific situations such as the impact of incomplete data (McArdle & Hamagami, 2001), or misspecification in the initial conditions (Ji & Chow, 2019).

The main goal of this study is to conduct a systematic evaluation of the extent to which the BLCS model is able to capture and distinguish the different sources of variability resulting from the interplay between developmental processes over time.

2. Method

We generated repeated measures for two latent processes (x and y), manipulating the following simulation conditions:

1. Sample size (two levels: n = 100, 200)
2. Number of repeated measures (five levels: T = 3, 5, 7, 10, 20)
3. Generating process (five models)

To compare the performance of BLCS models, we simulated 100 data sets for each of 50 conditions, formed by the combination of the previous conditions. Data sets were generated through five models with specific parameter combinations that represent realistic trajectories based on previous empirical studies. More details about the specification of these models are available from the authors upon request.

The baseline model (M1) includes two entirely independent processes. With this model, we want to evaluate the ability of the BLCS to detect the independence of the studied processes.

The (co)variances model (M2) includes variances and covariances for the additive components and the covariance between scores at time t = 0. With this model, we want to assess the constant amount of intra-individual change and the initial scores are correlated.

In the couplings model (M3), the two coupling effects γ_x and γ_y are added to the parameters of the baseline model. Through this model, we want to study the ability of the BLCS model to capture the generating parameters when additive components are individual-invariant, but time-lagged relations exist.

The dynamic error model (M4) includes dynamic error variances and covariances to the parameters of the baseline model. This model allows testing the recovery of stochasticity at the latent level, with entirely independent processes.

Finally, the complete model (M5) includes all the generating parameters. With this model, we want to study the ability of the BLCS model to capture and distinguish different sources of variability when covariances, coupling effects, and dynamic errors are all part of the true latent trajectories.

3. Results

We evaluated the rate of improper solutions, accuracy, efficiency, and 95%CI coverage for all conditions. In BLCS applications, the focus is usually on the time-lagged dynamics (i.e., self-feedback and cross-lagged effects). Because of this, here we focus on the estimation relative bias of self-feedbacks and couplings. They are depicted in Figure 2. The results for the remaining parameters are available from the authors upon request.

3.1. Number of repeated measures

The most relevant factor was the number of repeated measures. BLCS models can adequately recover all the parameters when seven or more repeated measures are used, with excellent rates of proper solutions (only 2 improper solutions across conditions with >7 measures), parameters accuracy (relative bias between 0.13-0.15), efficiency, and coverage (above 85% in all conditions). On the contrary, with three repeated measures, BLCS models yielded severely biased and inefficient estimates. With five repeated measures, the amount of bias in the parameter estimates decreased dramatically, leading to acceptable bias under many conditions (ranging from -.31 to .52). However, some parameters were still too inefficient. This finding has important implications in the design of longitudinal studies: in some conditions involving samples ≤ 200 cases, using a BLCS model requires at least seven measurement occasions to yield reliable results. With samples >200 cases, five repeated measures could be enough when initial scores and additive components are correlated.

3.2. Sample size

Regarding sample size, our results indicate that increasing the number of cases from 100 to 200 improves the recovery of the parameter estimates in terms of accuracy, efficiency, and coverage. However, this improvement is only meaningful with three and five repeated

measures. With seven or more repeated measures, the effects of including more cases are minimal. The most relevant finding here is that including one more measurement occasion improves the recovery of the parameter estimates to a much greater extent than including 100 more cases. For example, the relative bias range -1.51-2.53 obtained with 3 measures and 100 cases was reduced to -.31-.52 when 5 measures where used. However, increasing the number of cases to 200 with 3 measures produced only slight improvements: range -.82-1.45.

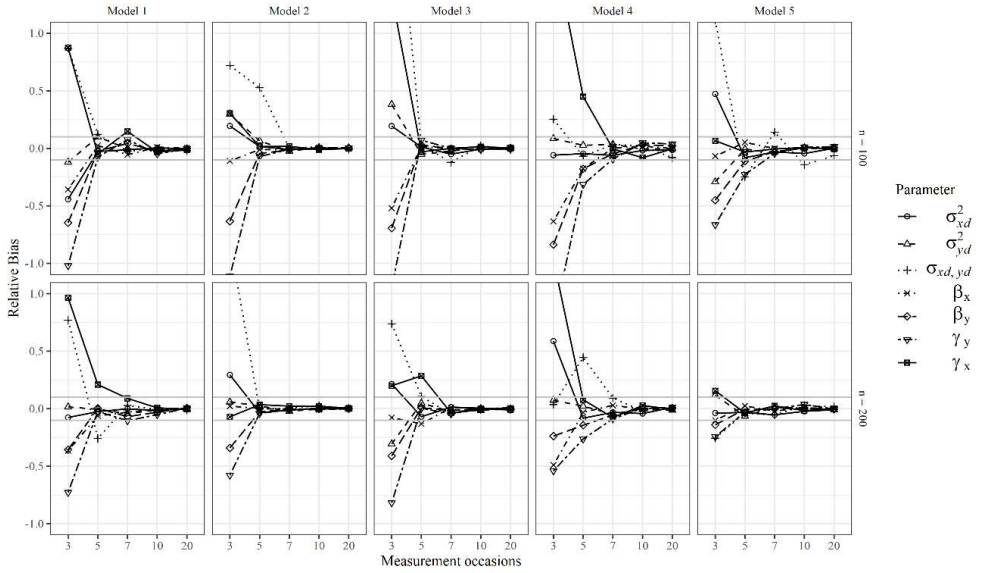

Figure 2. Relative bias for the dynamic error variances and covariances ($\sigma^2_{xd}, \sigma^2_{yd}, \sigma_{xd,yd}$), self-feedbacks ($\beta_x, \beta_y$), and couplings ($\gamma_x, \gamma_y$)

3.3. Generating process

The differences between the five sets of generating parameters become insignificant when seven or more repeated measures are used, yielding excellent results. Because of this, the following interpretations will be limited to conditions with three and five repeated measures.

On one hand, the (co)variances and complete models (M2 and M5) consistently led to more accurate and efficient estimates than the rest of the models. Acceptable results were found for them even with five repeated measures in samples of 200 cases. On the other hand, models M1, M3, and M4 led to poorer performance, probably due to the absence of between-individual variability in the additive components of both *x* and *y*. The couplings and variances of the additive components displayed the highest bias and variability. This is relevant since it indicates that the BLCS model will detect a significant amount of variability in the additive components, as well as interrelation between processes over time, even if those do not exist in the true underlying processes (M1 and M4). This poor performance is more remarkable

when dynamic error is present in the generating process (M4), leading to the highest rates of bias and variability in the parameter estimates, especially for the couplings. That is, the BLCS model struggles to capture and distinguish the different sources of variability when the two processes are independent, and the presence of dynamic error makes the recovery of the parameters even more difficult.

4. Conclusions

Based on the results from our extensive monte-carlo study, we offer the following recommendations to applied researchers who want to apply BLCS models in their longitudinal studies:

a) If a sample size of at least 100 cases is available, including more repeated measure is always preferable over including additional cases.
b) If seven repeated measures are available, including more measures will not significantly improve the quality of the parameter estimates.
c) If less than seven measures are available, self-feedbacks and couplings could be strongly biased when the processes under study are independent or dynamic error is present. If the researcher suspects that this may be the case with their data, we recommend including additional measures. If that is not possible, model estimates must be interpreted with caution.

We hope that these findings will inform the design of longitudinal studies, and provide insight about the risks of using BLCS models under adverse conditions of sample size and measurement occasions.

References

Bodenmann, G., Hilpert, P., Nussbeck, F. W., & Bradbury, T. N. (2014). Enhancement of couples' communication and dyadic coping by a self-directed approach: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *82*(4), 580-591. https://doi.org/10.1037/a0036356

Estrada, E., Sbarra, D. A., & Ferrer, E. (2020). Models for Dyadic Data. In A. G. C. Wright & M. N. Hallquist (Eds.), *The Cambridge Handbook of Research Methods in Clinical Psychology* (pp. 350–368). Cambridge University Press.

Ferrer, E., & McArdle, J. J. (2010). Longitudinal Modeling of Developmental Changes in Psychological Research. *Current Directions in Psychological Science*, *19*(3), 149-154. https://doi.org/10.1177/0963721410370300

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2007). Age changes in processing speed as a leading indicator of cognitive aging. *Psychology and Aging*, *22*(3), 558-568. https://doi.org/10.1037/0882-7974.22.3.558

Ghisletta, P., & Lindenberger, U. (2003). Age-Based Structural Dynamics Between Perceptual Speed and Knowledge in the Berlin Aging Study: Direct Evidence for Ability Dedifferentiation in Old Age. *Psychology and Aging*, *18*(4), 696-713. https://doi.org/10.1037/0882-7974.18.4.696

Ji, L., & Chow, S.-M. (2019). Methodological Issues and Extensions to the Latent Difference Score Framework 1. En E. Ferrer, S. M. Boker, & K. J. Grimm (Eds.), *Longitudinal Multivariate Psychology* (1.ª ed., pp. 9-37). Routledge. https://doi.org/10.4324/9781315160542-2

Kertes, D. A., Gunnar, M. R., Madsen, N. J., & Long, J. D. (2008). Early deprivation and home basal cortisol levels: A study of internationally adopted children. *Development and Psychopathology*, *20*(2), 473-491. https://doi.org/10.1017/S0954579408000230

King, L. A., King, D. W., McArdle, J. J., Saxe, G. N., Doron-LaMarca, S., & Orazem, R. J. (2006). Latent difference score approach to longitudinal trauma research. *Journal of Traumatic Stress*, *19*(6), 771-785. https://doi.org/10.1002/jts.20188

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. En R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling, present and future: A festschrift in honor of Karl Jöreskog* (pp. 7-46). Scientific Software International.

McArdle, J. J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology*, *60*(1), 577-605. https://doi.org/10.1146/annurev.psych.60.110707.163612

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637-657. https://doi.org/10.1037/met0000210

Identification of child self-regulation trajectories through growth mixture modelling

Alejandro de la Torre-Luque,[1,2] Andres Pemau,[1] Patricia Diaz-Carracedo,[1] and Cecilia A. Essau[3]

[1] Universidad Complutense de Madrid, Spain.

[2] Centre for Biomedical Research in Mental Health (CIBERSAM), Spain.

[3] University of Roehampton, UK.

Abstract

Purpose. This study aimed to illustrate the application of Growth mixture modelling (GMM) in the identification of self-regulation (SR) trajectories (emotional and behavioural SR) in childhood. Moreover, it intended to study how SR trajectories may be related to internalising symptoms and conduct problems in adolescence. **Methods.** Data from 13853 British participants (51.24% boys) from the Millennium Cohort Study were used. Participants' parents completed the Child Social Behaviour Questionnaire when participants aged 3, 5 and 7; and the Strength and Difficulties Questionnaire (SDQ) at participant's age 14. **Results.** Six trajectories of emotional SR (two of them, comprising 45.95% of participants, were considered risk trajectories due to elevated/increasing dysregulation) were identified as well as two trajectories of behavioural SR (a normative one and an at-risk trajectory, comprising 6% of participants). Participants showing a trajectory of child emotional dysregulation were featured by higher risk of emotional symptoms and conduct problems in adolescence. Conversely, normative behavioural regulation trajectory predicted adolescent emotional symptoms due to its overlap with emotion dysregulation. **Conclusion.** GMM seems to be useful to study the developmental course of SR in childhood, and to shed light into developmental mechanisms of adolescent mental disorder emergence.

Keywords: Growth mixture modelling; Longitudinal data analysis; Psychiatric Epidemiology; Child self-regulation; Adolescent mental health.

Funding: The Millennium Cohort Study was funded by the Economic and Social Research Council through grants awarded to the Centre for Longitudinal Studies (CLS).

***** *Corresponding author:* af.delatorre@ucm.es

1. State of the art

Self-regulation (SR) involves unfolding varying skills (e.g., cognitive abilities, behavioural tendencies) to deliberately modulate behavior, emotion and cognition towards adaptive adjustment (McClelland, Cameron, & Tominey, 2010). More concretely, SR abilities allow coping with situations featured by high emotional loading (the so-called hot self-regulation skills). In addition, SR is also involved in managing emotionally neutral situations, by unfolding skills (i.e., cool SR skills) such as inhibitory control or cognitive flexibility (Denham et al., 2012). Childhood constitutes a critical period for the acquisition of SR skills. In this regard, SR skills evolve in childhood as a result of an interplay between genetic predisposition, environmental factors and maturational processes (Lévesque et al., 2004; Rademacher & Koglin, 2019). Some studies have supported that adaptive SR increases with age from early childhood to adolescence (Gullone, Hughes, King, & Tonge, 2010; Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016). In contrast, maladaptive SR may rapidly evolve in early childhood but their growth rate tends to become decelerated from then on (Winsper & Wolke, 2014).

Cross-sectional and longitudinal studies have shown that maladaptive SR may be involved in mental disorder emergence (i.e., depression, anxiety, drug use disorders) as well as low academic achievement in childhood but also in adolescence (Edossa, Schroeders, Weinert, & Artelt, 2018; Eisenberg et al., 2017; Reinke, Herman, Day, & Connor, 2017). Fortunately, most children show a normative course featured by an increasing unfolding of adaptive skills to self-regulate over time, and lower levels of maladaptive skill unfolding. Even though, some risk trajectories have been identified regarding the emotional and behavioural domains. In this regard, it has been supported that some children may show a pattern of heightened emotional reactivity and deficient emotion regulation skills from childhood to adolescence (Cracco, Goossens, & Braet, 2017). On the other hand, risk trajectories of behavioural SR (featured by chronic disruptive behaviour, for instance) may increase the probability to persistently engage in violent interactions with others, drug abuse and delinquency throughout adolescence (Broidy et al., 2003; Pingault et al., 2013).

The identification of risk trajectories of maladaptive SR is of particular interest in order to feature the key mechanisms that may lead to impaired mental health in adolescence. In addition, the identification of SR trajectories may guide the development of preventive interventions to potentiate protective factors towards mental health promotion in the community. This study aimed to illustrate the use of growth mixture modelling (GMM) as a useful tool for the identification of risk trajectories of SR (emotional and behavioural SR) in childhood. Moreover, it intended to study how SR trajectories may be related to internalising symptoms and conduct problems in adolescence.

2. Method

Data from 13853 British toddlers (51.24% boys; mean age at baseline[1] = 3.13 years, *SD* = 0.20) from the Millennium Cohort Study (MCS; Connelly & Platt, 2014) were used. All the participants were born between 2000 and 2002 and lived in the United Kingdom at age 9 months (enrolment point).

The Millennium Cohort Study is a nationally-representative birth cohort study focused on: a) analysing patterns of physical and mental health development among toddlers born in the XXIst century (millennials); and b) identifying potential socioeconomic and health-related protective and risk factors. A stratified clustering strategy was followed in the MCS to ensure adequate representation of ethnic minorities. The MCS collects data from parents, teachers and cohort members. So far, data from seven survey waves have been collected in the MCS.

Due to the study purpose, we used data from parents' reports at MCS sweep 2 (cohort member age, *m* = 3.13, *sd* = 0.20), sweep 3 (cohort member age, *m* = 5.22, *sd* = 0.25) and sweep 4 (cohort member age, *m* = 7.23, *sd* = 0.25). More concretely, data from parents' responses to the Child Social Behaviour Questionnaire (CSBQ) were used. In addition, we used data from parents' responses to the Strength and Difficulties Questionnaire (SDQ) at cohort member's age 14 (MCS wave 6; cohort member age, *m* = 13.77, *sd* = 0.45). Finally, some other factors collected at wave 6 were considered in this study: a) from parents/legal guards (marital status, mental health status); b) environmental data (urbanicity of residence, income level); and c) cohort member data (sex, age, ethnicity, body mass index, physical health status, physical activity level, being in a romanting relationship, relationship with parents, feeling lonely, cognitive functioning).

In terms of analytical strategy, growth mixture modelling (GMM; Ram & Grimm, 2009) was used to identify heterogeneous trajectories that underlie an overall course. Bearing in mind the study goals, the latent process to be studied were emotion regulation and behaviour regulation across childhood, between age 3 to 8 years old (note that the CSBQ provides scores for two self-regulation factors: Emotion dysregulation and Behaviour regulation factors). GMM (as a person-centred approach) relaxes the assumption of a unitary course of regulation facets for all children. As a result, subject-specific variability can be better captured by clustering individuals with similar courses of development into a same group (class). In addition, it is assumed that person-specific effects (random effects) are collapsed into the mixture component (class-specific effects). Model estimation relied on robust maximum likelihood and full information methods (this enabled the depiction of individual-specific trajectories even when intermittent missing data were present). We estimated GMM models separately for emotion dysregulation (ED) and behaviour regulation (BR) processes. Age was

[1] Our baseline corresponds to the MCS wave 2

used as a time factor. Parameters estimated for each trajectory (intercept and time slope) class were considered in comparison to those from the first class identified (taken as a reference class). Covariates were not included in model estimation due to the increased probability of class overestimation (Vermunt, 2010). Under a model comparison tradition, solutions with increasing trajectory classes were tested, as well as those considering linear and quadratic effects on the fixed (overall course effects) and class-specific components (trajectory-specific effects). Criteria to select the model with the optimal class enumeration were: low sample-adjusted Bayesian information criterion (SABIC) and Akaike information criterion (AIC), mean of posterior probabilities to belong to each class higher than .70; and meaningful proportion of participants within each class (5%).

Multilevel regression was used to study how both ED and BR trajectory class memberships may predict emotional symptoms and conduct problem difficulties (SDQ subscales) at age 14. Covariates included were the sweep 6 factors (parent-related, environment and cohort member factors). The country (with factors: the UK countries) was the multilevel factor. The AIC was estimated to assess whether the model with covariates fitted better than the unconstrained model. The adjusted R^2 was used as an effect size estimate. The odds ratio (*OR*) was used as a predictor loading.

All the analyses were conducted by means of R x64 3.0.1 (lcmm and psych packages).

3. Results

Regarding the ED course, we found that the GMM model that fitted better to data was the 6-class model depicting a quadratic effects of time (SABIC = 24721.33, AIC = 24612.37; mean of posterior probabilities for each class = .87 - .97). The first class (U-shaped class; 15.25% of participants), was featured by decreasing ED with early ages, leveling off at age 6 (linear effect of age with slope, B = -0.67, p < .01; quadratic slope, B = 0.23, p < .01), tended to increase onwards (Figure 1). The second class identified (decreasing dysregulation class) comprised 22.51% of participants and featured by a decreasing trajectory of emotional dysregulation over time (linear effect of age with slope, B = -5.14, p < .01; quadratic slope, B = 0.93, p < .01). Another class (increasing dysregulation class) identified (10.94% of participants) was featured by rising emotional dysregulation over time (linear effect of age with slope, B = 0.41, p < .01; quadratic slope, B = -0.15, p < .01). The fourth identified class (heightened dysregulation class; 10.58% of participants) showed persistently elevated ED over time with slight increase in levels from age 6 from then on (linear effect of age, B = -1.59, p < .01; quadratic slope, B = 0.55, p < .01). ED course of participants from the fifth class (normative class; 32.34% of participants) showed minimal levels of ED up to age 6, with a slightly increasing pattern hereinafter (linear effect of age, B = -0.50, p < .01; quadratic slope, B = 0.17, p < .01). Finally, ED trajectory of the sixth class identified (inverted U-

shaped class; 8.39% of participants) was featured by increasing levels of ED up to age 6 and subsequent decrease with age (linear effect of age with slope, $B = 4.21$, $p < .01$; quadratic slope, $B = -0.60$, $p < .01$).

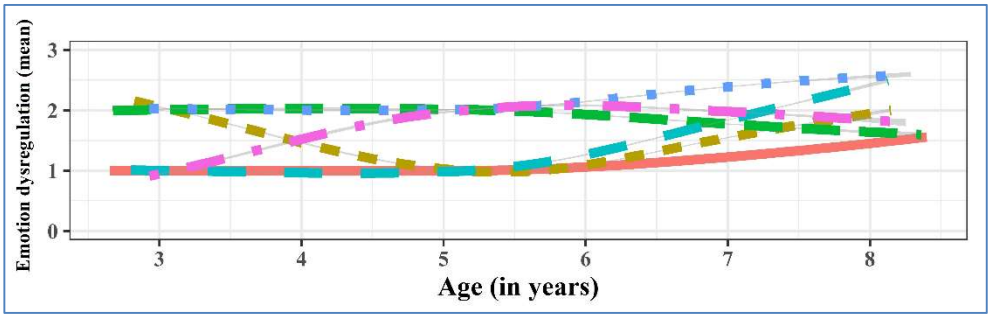

Figure 1. Trajectories of emotional dysregulation across trajectory classes.
Trajectory classes: U-shaped class = Olive green line. Decreasing dysregulation class = Green line. Increasing dysregulation class = Sky blue line. Heightened dysregulation class = Dotted blue line. Normative class = Solid pink line. Inverted U-shaped class = Dashdot violet line.

In terms of BR trajectory identification (see Figure 2), the 2-class GMM model, under a quadratic effects of age, fitted better to data (SABIC = 23802.21, AIC = 23762.98; mean of posterior probabilities for each class = .74 - .94; random-effects intercept = 0.59).

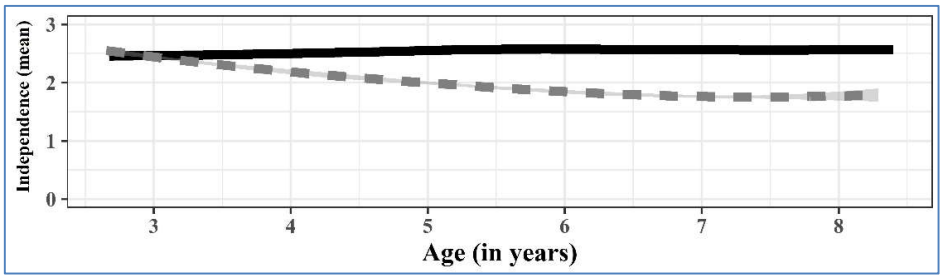

Figure 2. Trajectories of behavioural regulation (independence) according to trajectory classes.
Trajectory classes: Decreasing class = Dashed grey line. Normative class = Solid dark line.

The trajectory of participants from the first class identified (decreasing independence class; 6.07% of participants) was featured by a decreasing BR pattern over time (linear effect of age with slope, $B = -0.72$, $p < .01$; quadratic slope, $B = 0.06$, $p < .01$). On the other hand, the second class identified (normative class) comprised 93.93% of participants. Trajectory of BR

was featured by a slightly increasing pattern of independence over time (linear effect of age with slope, $B = 0.30$, $p < .01$; quadratic slope, $B = -0.04$, $p < .01$).

Multilevel regression modelling revealed that the covariates studied were important to explain the development of mental health difficulties at age 14 (for emotional symptoms: unconstrained model AIC = 67244.26, model with covariate AIC = 49974.07; for conduct problems: unconstrained model AIC = 56288.72, model with covariate AIC = 39852.04). The models with covariates explained a significant proportion of outcome variance, for both emotional symptoms ($R^2_{adj} = .27$) and conduct problems ($R^2_{adj} = .32$). The Table 1 displays the coefficients of ED class membership (in comparison to the normative ED class) and those of the decreasing independence class (in comparison to the normative BR class) to explain the emotional symptoms and conduct problems at age 14.

Table 1. Self-regulation trajectory classes and their predictive loading on outcomes in adolescence.

	Emotional symptoms			**Conduct problems**		
	OR	***CI*$_{95}$**	***Z***	***OR***	***CI*$_{95}$**	***Z***
Emotion dysregulation[1]						
U-shaped	1.05	1.04, 1.07	6.82	1.14	1.12, 1.15	15.69
Decreasing dysregulation	1.02	1.01, 1.04	2.77	1.07	1.05, 1.09	7.28
Increasing dysregulation	1.04	1.02, 1.06	4.33	1.09	1.07, 1.12	9.45
Heightened dysregulation	1.13	1.11, 1.15	15.40	1.26	1.24, 1.28	26.50
Inverted U-shaped	1.02	1.00, 1.03	1.86^{ns}	1.05	1.03, 1.06	4.81
Behavioural regulation[1]						
Decreasing independence	0.96	0.95, 0.98	-5.12	0.91	0.77, 1.08	-1.06^{ns}

Note. The regression models are adjusted for child, parents and environmental factors. Outcomes are subscales from the Strength and Difficulties Questionnaire completed when cohort member was 14 years.
OR = Odds ratio. *CI*$_{95}$ = 95% confidence interval of the OR. *Z* = Contrast test. All *OR* were significant ($p < .01$), except those with ^{ns} ($p > .05$). [1] Ref. class = normative class.

Being a member of each ED classes showed a significant loading to explain both the development of emotional symptoms and conduct problems in adolescence (in comparison to normative class), except the inverted U-shaped class when considering the development of emotional symptoms. Participants from the risk trajectory classes (i.e., U-shaped class, increasing class, and heightened dysregulation class) showed the highest risk of symptom/problem development, but particularly those from the heightened dysregulation class. On the other hand, being a member of the normative BR class only predicted the development of emotional symptoms in adolescence.

4. Conclusions

This study provided some interesting evidence on the relevance of addressing inter-individual variability in the depiction of self-regulation development in childhood. First, we found that heterogeneous trajectories depicted better the course of both ED and BR processes than unitary courses in childhood. Second, trajectory class identification univocally helps to feature risk endophenotypes. In this regard, we observed that the risk trajectories of ED (i.e., U-shaped class, increasing class, and heightened dysregulation class) were related to increased risk of showing both emotional symptoms and conduct problems in adolescence. Note that 36% of participants were classified into the risk ED trajectories. Moreover, showing a heightened ED trajectory in childhood put adolescents at the highest risk for mental health development. These results are free from the influence of multiple potential confounders, coming from internal and external environment. Our results go in line with those from other studies on developmental psychopathology (Eisenberg et al., 2017; Pingault et al., 2013). Surprinsingly, the relationship between BR trajectory and development of adolescent mental health problems were opposite than expected (i.e., a higher risk of emotional symptom development was of observed in terms of being a member of the normative BR class). This finding should be considered taking into account the overlap with ED risk trajectories (more than 68% of adolescents in the normative BR class showed a risk trajectory of ED). This may point to emotional regulation should be a prominent facet of SR with higher impact on mental health at short and long terms, in comparison to other SR facets.

In conclusion, GMM allows for great flexibility in featuring patterns of change and continuity of developmental processes. A latent or unobserved variable (trajectory class) accounting for inter-individual variability can be described, with a particular focus on depicting groups with homogeneous courses. Finally, the use of GMM, as a person-centred method, may definitively contribute to the advancement of medicine precision, by means of risk trajectory characterisation.

References

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., … Vitaro, F. (2003). Developmental Trajectories of Childhood Disruptive Behaviors and Adolescent Delinquency: A Six-Site, Cross-National Study. *Developmental Psychology*, *39*(2), 222–245. https://doi.org/10.1037/0012-1649.39.2.222

Connelly, R., & Platt, L. (2014). Cohort profile: UK Millennium Cohort Study (mcs). *International Journal of Epidemiology*, *43*(6), 1719–1725. https://doi.org/10.1093/ije/dyu001

Cracco, E., Goossens, L., & Braet, C. (2017). Emotion regulation across childhood and adolescence: evidence for a maladaptive shift in adolescence. *European Child and Adolescent Psychiatry*, *26*(8), 909–921. https://doi.org/10.1007/s00787-017-0952-8

Denham, S. A., Bassett, H. H., Way, E., Mincic, M., Zinsser, K., & Graling, K. (2012). Preschoolers' emotion knowledge: Self-regulatory foundations, and predictions of early

school success. *Cognition and Emotion*, *26*(4), 667–679. https://doi.org/10.1080/02699931.2011.602049

Edossa, A. K., Schroeders, U., Weinert, S., & Artelt, C. (2018). The development of emotional and behavioral self-regulation and their effects on academic achievement in childhood. *International Journal of Behavioral Development*, *42*(2), 192–202. https://doi.org/10.1177/0165025416687412

Eisenberg, N., Hernández, M. M., Spinrad, T. L., Eisenberg, N., Hernández, M. M., & Spinrad, T. L. (2017). The Relation of Self-Regulation to Children's Externalizing and Internalizing Problems. In *Emotion Regulation and Psychopathology in Children and Adolescents* (pp. 18–42). https://doi.org/10.1093/med:psych/9780198765844.003.0002

Gullone, E., Hughes, E. K., King, N. J., & Tonge, B. (2010). The normative development of emotion regulation strategy use in children and adolescents: A 2-year follow-up study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *51*(5), 567–574. https://doi.org/10.1111/j.1469-7610.2009.02183.x

Lévesque, J., Joanette, Y., Mensour, B., Beaudoin, G., Leroux, J. M., Bourgouin, P., & Beauregard, M. (2004). Neural basis of emotional self-regulation in childhood. *Neuroscience*, *129*(2), 361–369. https://doi.org/10.1016/j.neuroscience.2004.07.032

McClelland, M. M., Cameron, C. E., & Tominey, S. L. (2010). Self-regulation : The integration of cognition and emotion. In R. Lerner & W. Overton (Eds.), *Handbook of life-span development (vol. 1: Cognition, biology and methods)* (pp. 509–553). Hoboken, NJ: Wiley & Sons.

Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology*, *52*(11), 1744–1762. https://doi.org/10.1037/dev0000159

Pingault, J. B., Côté, S. M., Galéra, C., Genolini, C., Falissard, B., Vitaro, F., & Tremblay, R. E. (2013). Childhood trajectories of inattention, hyperactivity and oppositional behaviors and prediction of substance abuse/dependence: A 15-year longitudinal population-based study. *Molecular Psychiatry*, *18*(7), 806–812. https://doi.org/10.1038/mp.2012.87

Rademacher, A., & Koglin, U. (2019). The concept of self-regulation and preschoolers' social-emotional development: a systematic review. *Early Child Development and Care*, *189*(14), 2299–2317. https://doi.org/10.1080/03004430.2018.1450251

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioural Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Reinke, W. M., Herman, K. C., Day, S. L., & Connor, C. M. D. (2017). Examining the Relations between Self-Regulation and Achievement in Third-Grade Students. *Assessment for Effective Intervention*, *42*(2), 97–109. https://doi.org/10.1177/1534508416670367

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*(4), 450–469. https://doi.org/10.1093/pan/mpq025

Winsper, C., & Wolke, D. (2014). Infant and toddler crying, sleeping and feeding problems and trajectories of dysregulated behavior across childhood. *Journal of Abnormal Child Psychology*, *42*(5), 831–843. https://doi.org/10.1007/s10802-013-9813-1