Available online at www.sciencedirect.com

SCIENCE DIRECT®

Learning and Individual Differences 16 (2006) 159–174

Learning and Individual Differences

www.elsevier.com/locate/lindif

Issues in longitudinal research on motivation☆

Reinoud D. Stoel [a,*], Jaap Roeleveld [a], Thea Peetsma [a], Godfried van den Wittenboer [a], Joop Hox [b]

[a] University of Amsterdam, Department of Education, P.O. Box 94208, NL-1090 GE Amsterdam, The Netherlands
[b] Utrecht University, The Netherlands

Received 28 January 2005; received in revised form 26 July 2005; accepted 27 July 2005

Abstract

This paper discusses two methodological issues regarding the analysis of longitudinal data using structural equation modeling that emerged during the reconsideration of the analysis of a recent study on the relationship between academic motivation and language achievement in elementary education [Stoel R.D., Peetsma, T.T.D. and Roeleveld, J. (2003). Relations between the development of school investment, self-confidence and language achievement in elementary education: a multivariate latent growth curve approach. Learning and individual differences, 13, 313–333]. The issues are related to the factorial structure of the repeatedly measured variables, and to the explanation of interindividual difference by means of covariates [see Stoel, R.D., Van den Wittenboer, G. and Hox, J.J. (2004a). Including time-invariant covariates in the latent growth curve model. Structural Equation Modeling, 11, 155–167, Stoel, R.D., Van den Wittenboer, G. and Hox, J.J. (2004b). Methodological issues in the application of the latent growth curve model. In K. van Montfort, H. Oud, and A. Satorra (Eds.). Recent developments on structural equation modeling: Theory and applications. (pp. 241–262). Amsterdam: Kluwer Academic Press. It is illustrated that standard modeling practices may sometimes lead to incorrect conclusions regarding the concepts under investigation, and that ideally alternative modeling possibilities should be considered in order to check the adequacy of the standard practice.

Keywords: Motivation; Structural equation modeling; Latent growth curve analysis

☆ Author note: This research was made possible by a grant from the Netherlands Organization for Scientific Research (NWO/PROO).
* Corresponding author. Tel.: +31 20 525 1529; fax: +31 20 525 1200.
E-mail address: r.d.stoel@uva.nl (R.D. Stoel).

1041-6080/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.lindif.2005.07.006

1. Introduction

In the last decade there has been an increasing amount of studies on academic motivation that adopted a longitudinal design. Longitudinal designs may provide important information for answering longstanding questions regarding change and growth of individuals on motivation. In order to answer such questions complex models, and techniques, have been developed, and these models and techniques are now becoming part of the standard tool box of many scholars. Examples are structural equation modeling and multilevel analysis of longitudinal data, latent class analysis, and (growth) mixture modeling. However, because of the complexity of these techniques, the application is often plagued by factors that may lead to incorrect estimates of the interesting parameters, and thus to possibly incorrect conclusions. Standard, but sometimes insufficient, modeling practice may have serious consequences for the substantive conclusions. This contribution discusses two methodological issues regarding the analysis of longitudinal data using structural equation modeling that emerged during the reconsideration of the analysis of a recent study on the relationship between academic motivation and language achievement in elementary education (Stoel, Peetsma & Roeleveld, 2003). The issues are related to the factorial structure of the repeatedly measured variables, and to the explanation of interindividual difference by means of covariates. A more formal and detailed treatment of these two issues is provided by Stoel, van den Wittenboer and Hox (2004a,b). The first purpose of this paper is to illustrate that the standard approaches can be easily adapted to overcome these inadequacies, and second to provide practical guidelines on how and when to do so. In the next sections we will first describe the data, and the sample and the variables that were measured, then we will provide a brief introduction into latent growth curve modeling, followed by an overview of the analysis strategy and the results of Stoel, Peetsma and Roeleveld, and successively the two issues will be discussed.

2. Latent growth curve modeling of motivation, school investment and language acquisition

The study of Stoel, Peetsma and Roeleveld (2003) was guided by the following main questions and expectations: (1) How do school investment, self-confidence and language achievements develop during elementary education (from kindergarten to secondary education)? An increase in language achievement is expected, and a decrease in school investment during elementary education. With respect to self-confidence, no expectation was formulated on the direction of development during elementary education. (2) Is the developmental process of language achievement in elementary education related to school investment and self-confidence? It is expected that the developmental trajectories of language achievement, school investment and self-confidence are mutually positively associated. With respect to school investment, this implies that the more positive the developments in achievement and self-confidence, the less the decrease in school investment. (3) To what extent is intelligence related to developmental trajectories in school investment, self-confidence and language achievements in elementary education? It is expected that intelligence accounts for a unique part of the variation in the developmental trajectories of language achievement.

In order to answer these questions data from the large PRIMA cohort project in the Netherlands were analyzed. These data consist of a subsample consisting of 2693 children in 214 elementary schools, measured at four consecutive points in time (every 2 years, from grade 2 at the age of 5 up to grade 8 at the age of 11 years). Measurements of Language achievement, School Investment and Self-confidence were available at each of the four time points, as well as a measure of intelligence. Language achievement was measured by four different tests at different ages. The individual scores were consequently transformed to one language ability scale, using OPLM-procedures from item-response theory (Verhelst, Glas & Verstralen, 1993; Vierke, 1995). Because of the young age of the children, no direct measurements of Self-confidence and School Investment were performed in the PRIMA project. As a proxy, indicators were used taken from an instrument which was filled out by the teachers. In order to have identical measurements over the years, only items that have been unchanged could be used. Both Self-confidence and School Investment were each measured by two indicators at the four points in time. Further information on the data and the sample is provided by Stoel, Peetsma and Roeleveld. Appendix A presents the estimated covariance matrix and means vector.

To describe the development of each of these three concepts latent growth curve (LGC) modeling (McArdle, 1986, 1988; Meredith & Tisak, 1990; Willett & Sayer, 1994) has been applied. LGC analysis is a special type of structural equation model that lets repeated measures of a given concept to be represented as a function of time and other measures (Willett & Sayer, 1994). A LGC analysis assumes an individual growth curve for each subject to account for the development over time, with the assumption that all subjects in a given population have developmental curves of the same functional form (e.g. all linear), but with possibly different growth parameters, i.e. the (initial) level, or intercept, and the growth rate, or slope. Regarding linear developmental curves, for example, individual differences may be due to heterogeneity in the (initial) level, as well as heterogeneity in the growth rate (or rate of change). So, subjects may differ in their level at the first measurement occasion and develop subsequently at different rates. In the current research the focus is on the intra- and interindividual variation of the three repeatedly measured variables (Language ability, Self-confidence and School investment), in particular with regard to the covariation across individuals with respect to the patterns of development on these variables (cf. MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997). A short introduction to LGC analysis is presented below. More detailed introductions to latent growth curve modeling are presented, for instance, by Willett and Sayer (1994), MacCallum et al. (1997), Muthén and Khoo (1998) and Duncan, Duncan, Strycker, Li and Alpert (1999).

As noted by Muthén and Khoo (1998) presenting the level and the growth rate as latent variables enables conventional structural equation modeling (SEM) software to analyze the growth curve models, and path diagrams can then be drawn to give a graphic presentation. Fig. 1 presents such a path diagram of a growth curve model for language acquisition. While the individual growth curves are not shown explicitly, the diagram shows the means and variances of the parameters describing the growth curves, and their correlation; lan_2 to lan_8 represent the consecutive measurements of the outcome variable language acquisition. The ellipses represent the latent variables, respectively *Level* and *Growth rate*, and contain the means (regression on the diamond) and variances of respectively, the (initial) level and the growth rate of the curves. These latent variables provide information on the interindividual differences in the developmental curves, and are allowed to correlate. Of primary interest is the growth rate, which corresponds to development over the subsequent years.

Fig. 1 actually is a confirmatory factor model, with the difference that the factor loadings are constrained to specific values. This way of presenting the models enables an interpretation of the latent factor as chronometric latent factors, instead of the usual psychometric latent factors (McArdle, 1989).

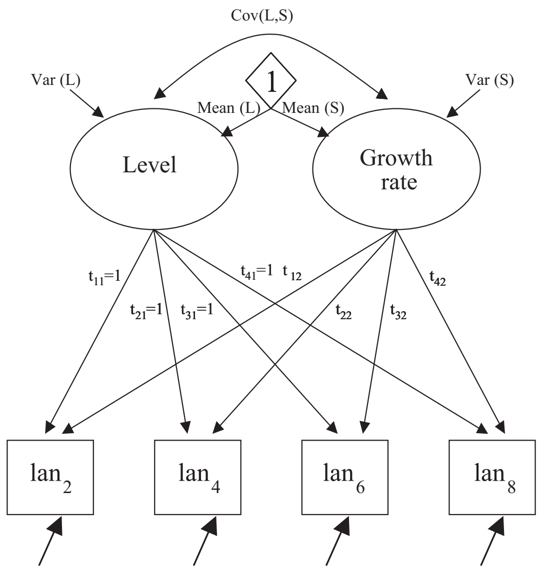

Fig. 1. Graphic representation of latent growth curve model.

The latent growth curve model can either be build on observed indicators, or on latent factors. So, if multiple indicators of a given concept are available the factorial structure of these measurements can be accounted for in the model. The latent growth curve model is then build on the first-order latent factors, instead of on the observed indicators. In other words, the common variation in the multiple indicators is accounted for by the first-order factors, while the second-order factors serve to explain the mean and covariance structure of the first-order factors. Illustrative examples of a LGC model with multiple indicators can be found in Garst, Frese and Molenaar (2000), and Hancock, Kuo, and Lawrence (2001). Thus, instead of analyzing the sum scores or item parcels, as is often the case, the observed indicators can be put directly into the analysis. In the present study the multiple indicators of the motivational variables, self-confidence and school investment, were explicitly incorporated in the model.

Incorporating the factorial structure explicitly into the model also allows for the test of an important assumption, i.e. the assumption of measurement invariance. This assumption cannot be tested if the growth curve model is build directly on the indicators. Measurement invariance (the invariance of the measurement parameters) ensures a comparable definition of the latent construct over time (Hancock, Kuo & Lawrence, 2001). Measurement invariance can exist to a certain degree. If some assumptions are violated the longitudinal factor model is said to have partial measurement invariance (Byrne Shavelson & Muthén, 1989; Pentz & Chou, 1994). Practically spoken, a test of the assumption of measurement invariance implies equality constraints on the factor loadings and intercepts of the repeatedly measured, and modeled, variables. One way of implementing the assumption of measurement invariance on the first-order latent factors in the latent growth curve model is presented in Fig. 2, which presents a growth curve model for the school investment. For purposes of scaling the factor loadings for all $inv1_t$ are fixed to 1; also the intercept for all $inv1_t$ is constrained to zero. Furthermore, the factor loadings and intercepts of $inv2_t$ are constrained to be equal over time. In this specific model $inv1_t$ is said to be the reference indicator since it is used to scale the first-order latent

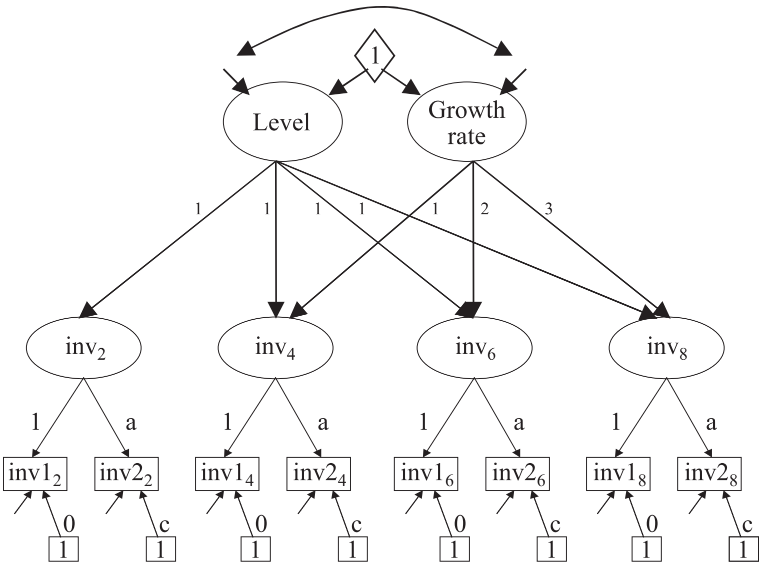

Fig. 2. Schematic presentation of full measurement invariance in the latent growth curve model for school investment. Note: Intercepts of indicators are conceptualized as regression on a constant equal to one (See Hancock et al. 2001). Only the relevant parameters are presented. Factor loadings for $inv1_t$ are fixed to 1.00 prior to estimation; factor loadings of $inv2_t$ are constrained to be equal (a); intercepts of $inv1_t$ are fixed to zero; intercepts of $inv2_t$ are constrained to be equal (c). The curved double-headed arrows represent correlations between the latent factors. [a]Factorloadings fixed to 1.00 prior to estimation.

factors of school investment. Alternatively also $inv2_t$ could have been used as the reference indicator, this would lead to a statistically equivalent model.

2.1. Short overview of the analysis strategy and the results of Stoel, Peetsma and Roeleveld (2003)

For the models reported in this article M*plus* was used (Muthén & Muthén, 1998) to fit the models to the data. Both means and covariances are analyzed simultaneously. So, unlike usual structural equation models, the factor means are not assumed to be zero. The chi-square measure of overall goodness of fit (χ^2), in combination with the root mean square error of approximation (RMSEA, Browne & Cudeck, 1992), will be used to evaluate the overall goodness of fit of the models. For the comparison of competing models the chi-square difference test ($\Delta\chi^2$), can be used. Our sample has a considerable amount of missing data, but most of these are missing 'by design'. These data can be considered to be missing completely at random (MCAR; Little & Rubin 1987). The other part of the missing data were not missing by design, but missing due to other causes. Although it is still common practice to use naive methods such as listwise, or pairwise deletion to deal with the missing data problem, these methods have been criticized extensively (Little & Rubin, 1987). In our analysis the Full Information Maximum Likelihood (FIML) estimation procedure of Muthén, Kaplan and Hollis (1987) was used, as implemented in the software package M*plus*. In constructing our final model, the strategy consisted of two parts: (1) testing the measurement models, and (2) combining the measurement models in a structural model (Anderson & Gerbing, 1988). The main rationale behind this strategy is that possible model misspecifications can be located more easily, and necessary additional tests can be performed more adequately.

First, longitudinal measurement models were fit for the concepts of school investment and self-confidence. These models form the basis for further analyses. The models of school investment and self-confidence were found to follow a similar pattern. For each model, the factor loadings of the indicators could be constrained to be equal over time, and the intercepts at all but the first occasion, besides the necessary constraints for identification purposes. Residuals of the same indicators were correlated across time up to two occasions, a so-called "partially banded" residual structure (see Vonesh & Chinchilli, 1997). The measurement models both have good overall goodness of fit measures. The chi-square measure of fit is non-significant ($\alpha = .01$) and the RMSEA <.05. Self-confidence has a fit of: $\chi^2(9) = 17.18$, $p = .05$; RMSEA = .018, and School investment: $\chi^2(9) = 8.97$, $p = .44$; RMSEA = .000. It is

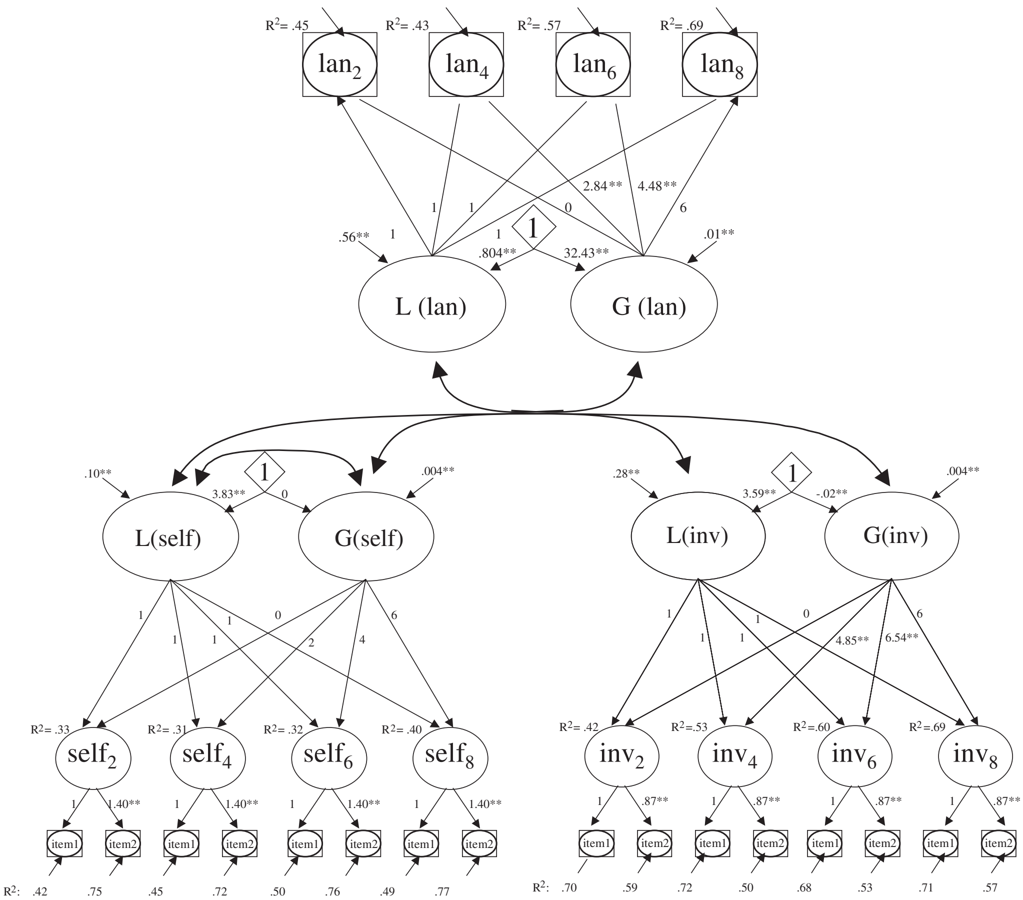

Fig. 3. Multivariate growth curve model of language ability, self-confidence and school investment. Note. lan=Language ability, self=Self-confidence, inv=School investment. From "Relations between the development of school investment, self-confidence and language achievement in elementary education: a multivariate latent growth curve approach." by R.D. Stoel, T.T.D. Peetsma and J. Roeleveld, 2003, Learning and Individual Differences, 13, p.324. Copyright by Elsevier Science Inc.

Table 1
Regression coefficients of the growth parameters on intelligence, and correlations between growth parameters in the multivariate growth curve model controlled for intelligence

	1.	2.	3.	4.	5.	6.
Growth parameters on intelligence	.54**	.28**	.09	.23**	.47**	.02
1. Level (language)	1					
2. Growth (language)	0	1				
3. Level (self-confidence)	.12**	−.15**	1			
4. Growth (self-confidence)	0	.35**	−.42*	1		
5. Level (school investment)	.18**	.06	.00	.14**	1	
6. Growth (school investment)	0	.27**	0	0	0	1

**$p<.01$. From "Relations between the development of school investment, self-confidence and language achievement in elementary education: a multivariate latent growth curve approach." by R.D. Stoel, T.T.D. Peetsma and J. Roeleveld, 2003, *Learning and Individual Differences*, 13, p.324. Copyright by Elsevier Science Inc.

concluded that partial measurement invariance has been found for both models. The intercepts of the indicators at the first measurement occasions for self-confidence and school investment could not be constrained to be equal to the intercepts at other occasions. Apparently the children interpreted the questions at the first measurement occasion in a slightly different manner compared to the other occasions.

Secondly, latent growth curve models were fit on the first-order factor structure for school investment and self-confidence, and for language acquisition on the observed scores. The results indicate a well-fitting linear growth curve model for the concept of self-confidence ($\chi^2(15)=27.27$, $p=.03$, RMSEA=.018), and a nonlinear growth model for both school investment ($\chi^2(13)=14.72$, $p=.33$, RMSEA=.007) and language ability ($\chi^2(4)=21.53$, $p=.00$, RMSEA=.040). The results of our study showed significant differences between children in their developmental curves in all three processes. The parameter estimates of the three models are presented jointly in Fig. 3. These results support the expectation of the decrease in school investment over the period of attendance at elementary school, except for the small increase at the end of the school period. Development in language ability, school investment and self-confidence were, as can be seen in Table 1, mutually positively associated. Children having a greater increase in their language ability also had a greater increase in their self-confidence, as well as a smaller decrease in their school investment. These results are in accordance with the second hypothesis. Furthermore, it was found that intelligence accounts for some of the individual differences in the development of language ability, school investment and self-confidence. The more intelligent the children, the higher is their level at the start, and the more positive is their development on all three concepts. Regarding the second point, it was found that intelligence accounts for some, but not all of the association between the developmental processes. The fact that the processes of school investment and self-confidence are relatively unrelated supports the argument that school investment, self-confidence and intelligence each may explain a different portion of the individual differences in the development of language ability. This result offers confirmation for the third hypothesis.

In this paper we will take a closer look at two aspects of the above described study. In the next section we will give attention to the measurement models for self-confidence and school investment and we will examine consequences of different modeling for our results. In section 3 we will discuss the modeling of intelligence as time-invariant covariate of the growth parameters.

3. Measurement invariance of multiple indicator concepts

The ability to model multiple indicators under a common factor is one of the advantages of SEM, as is shown in the extensive literature dealing with issues concerning common factors and structural equation models in general (e.g. Bollen, 1989). However, incorporating the factorial structure explicitly in to the model, asks for a choice concerning the scaling of the latent variable structure. As noted before, a common approach to scale the first-order latent variable structure in a latent growth curve model is to identify a reference variable, preferably the same at each occasion, and to constrain its factor loading to 1 and its intercept to zero (Hancock, Kuo & Lawrence, 2001; Oort, 2001; Sayer & Cumsille, 2001). Stoel, Van den Wittenboer and Hox (2004b) demonstrate that under full measurement invariance the choice of the reference indicator does not affect the parameter estimates and the model fit and the same substantive conclusions will be drawn.

Partial measurement invariance is not as strict as full measurement invariance in that a few violations are tolerated. That is, the factor loadings (a), and or the indicator's intercepts (c) do not have to be of the same value, and need not necessarily constrained to be equal for the full time period. Following the arguments for full measurement invariance, it may be expected that the model is also insensitive to the scaling of the latent variables. This is not always true, however, if a different reference indicator is used different parameter estimates could emerge. In other words, the latent growth curve model under partial measurement invariance is not invariant under a different scaling of the mean structure of the latent variable by using a reference indicator (see Stoel, Van den Wittenboer & Hox, 2004b). We will now explore what the consequences are of the choice for the specific reference indicators for the parameter estimates and model fit of the growth curve models for school investment and self-confidence.

The analyses of the longitudinal measurement models for school investment and self-confidence started out with the unconstrained longitudinal factor models. Subsequently, the constraints for measurement invariance were imposed, and their tenability was assessed. Fig. 4 gives a schematic presentation of the longitudinal factor models for self-confidence, but a similar figure can be drawn for school investment. In this figure $self1_t$ represents the first indicator of self-confidence, and $self2_t$ represents the second indicator. In the case of school investment these indicators would be, respectively,

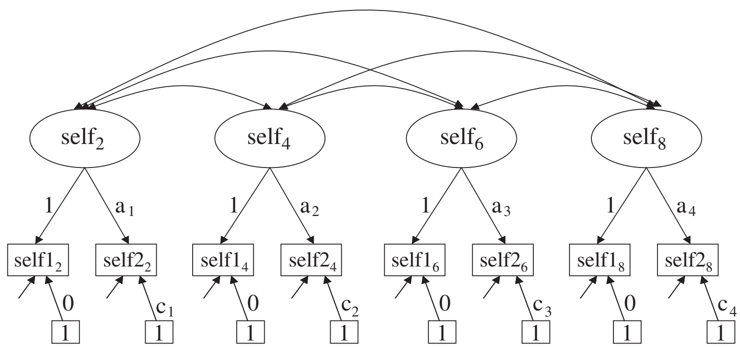

Fig. 4. Schematic presentation of an unconstrained longitudinal factor model for self-confidence. Note. Residuals are correlated across time up to two occasions.

inv1$_t$ and *inv2*$_t$. The necessary constraints for identification purposes were implemented on one of the two items, but the remaining factor loadings (a_t) and indicator intercepts (c_t) were freely estimated. Table 2 presents the fit statistics of the nested sequences of longitudinal measurement models for both concepts, in which the first indicator (respectively *self1*$_t$ and *inv1t*) at each occasion was used to scale both the covariance and mean structure.

Summarizing the results in Table 2, for both self-confidence and school investment equal factor loadings of the second indicator across time did not lead to a significant decrease in model fit. However, a subsequent equality constraint on the intercepts gave a significant deterioration in model fit. The modification indices showed that, for both models, the misfit was due to the intercepts at the first measurement occasion. These modification indices for the intercept of the indicators at the first occasion for self-confidence and school investment were, respectively, 14.96 and 83.41. Though it did not matter for the model fit which constraint was relaxed (*self1*$_t$ or *self2*$_t$ vs. *inv1*$_t$ or *inv2*$_t$), it was decided to remove the equality constraint on the second indicator (*self2*$_t$ vs. *inv2*$_t$) in order not to change the identification scheme. The partially constrained longitudinal measurement model provided a good fit to the data, as was already reported before.

The first question is now what would have happened if the second indicator (*self2*$_t$ vs. *inv2*$_t$) was used for identification purposes instead of the first indicator. The results concerning the model fit were equivalent for all tested models. Therefore, only the fit measures of the final models are reported (Model 1.5 and 2.5). Table 2 shows that a different reference indicator did not lead to any

Table 2
Fit results for the longitudinal measurement models

Model		*df*	χ^2 (*p*)	RMSEA
Self-confidence				
1.1	No restrictions	4	7.22 (.12)	.017 (.000–.037)
1.2	Equal factor loadings	7	16.84 (.02)	.023 (.009–.037)
	Difference 1.2 and 1.1	7 − 4 = 3	9.62 (.02)	
1.3	Equal intercepts	10	32.31 (.00)	.029 (.018–.040)
	Difference 1.3 and 1.2	10 − 7 = 3	22.65 (.00)	
1.4	Partially equal intercepts	9	17.18 (.05)	.018 (.002–.032)
	Difference 1.4 and 1.2	9 − 7 = 2	.30 (.86)	
1.5	Different scaling partially equal intercepts	9	17.18 (.05)	.018 (.002–.032)
School investment				
2.1	No restrictions	4	5.11 (.28)	.010 (.000–.32)
2.2	Equal factor loadings	7	7.99 (.33)	.007 (.000–.026)
	Difference 2.2 and 2.1	3	2.88 (.41)	
2.3	Equal intercepts	10	93.48 (.00)	.056 (.046–.067)
	Difference 2.3 and 2.2	3	88.37(.00)	
2.4	Partially equal intercepts	9	8.97 (.44)	.000 (.000–.022)
	Difference 2.4 and 2.2	2	.98 (.61)	
2.5	Different scaling partially equal intercepts	9	8.97 (.44)	.000 (.000–.022)

N = 2660; ** p < .01; RMSEA values in parentheses denote 90% confidence intervals; Models 1.4 and 2.4 are estimated without an equality constraint on the intercepts of the indicator at the first occasion.

Table 3
Goodness of fit measures of the univariate growth curve models

Model		*df*	$\chi^2(p)$	RMSEA
First indicator as reference				
Self-confidence	Model 1.2	15	27.27 (.03)	.018 (.006–.028)
School investment	Model 2.2	13	14.72 (.33)	.007 (.000–.021)
Second indicator as reference				
Self-confidence		15	21.31 (.12)	.013 (.000–.024)
School investment		13	14.38 (.35)	.006 (.000–.021)

The 90% confidence interval for RMSEA is given in brackets.

difference in model fit. Both Model 1.5 and 2.5 have equivalent fit measures as, respectively, Model 1.4 and 2.4. Apparently, there are no consequences for the model fit of the choice for a specific reference indicator under partially equal indicator intercepts if the structural part of the model is unconstrained.

Next, the same growth structure is imposed on the structural part of the model as was found by Stoel, Peetsma and Roeleveld (2003). Recall that a linear growth curve model was found with a zero mean growth for self-confidence, and a nonlinear growth curve model with a zero correlation between the level and growth rate for school investment. Table 3 presents the results. The first rows present the fit measures of the models as analyzed and reported in Table 1 of Stoel, Peetsma and Roeleveld, using the first indicator ($self1_t$ vs. $inv1_t$) as the reference, the last two rows present the growth curve models were the reference indicator has been changed.

As expected from the results of Stoel, Van den Wittenboer and Hox (2004b), the models with a different reference indicator are not exactly equal under partial measurement invariance. However, the differences are, in this case, not so large that they would lead to different substantive conclusions. For both models the model fit gets slightly better if the second indicator is used as the reference indicator; and the parameter estimates show minor changes. As a final check, the multivariate growth curve model was analyzed again using the second indicator of self-confidence and school investment as the reference for scaling. The fit measures of this model were $\chi^2(180)=341.35$, $p=.00$, RMSEA=.017, approximately the same as the final model of Stoel, Peetsma and Roeleveld (2003). It may be concluded that in this particular situation the impact of partial measurement invariance was not very important. Though model fit improved slightly, the parameter estimates and substantive conclusions remained the same.

4. Modeling time-invariant covariates as predictors of the growth parameters

In this section we will discuss the way the time-invariant covariate 'Intelligence' was incorporated in the multivariate growth curve model presented before. This was done using the growth predictor model (see Stoel, Van den Wittenboer & Hox, 2004a), in which the growth parameters of the three processes, initial level and growth rate, were regressed on intelligence. The results of Stoel, Van den Wittenboer and Hox, however, indicate that this might not always be the best approach. That is, if the assumption of full mediation is violated, the results of the growth predictor model may be biased, and substantive conclusions may be incorrect. The assumption of full mediation states that the effect of the time-invariant covariate on the observed indicators is fully mediated by the latent growth parameters. In this section, results of additional analyses are presented in which intelligence is modeled using the direct effect model of Stoel, Van den Wittenboer and Hox. The main question here is whether the growth predictor model is justified by testing if the assumption of full mediation is violated for these data, and if so, what the implications are for the reported results and conclusions. The assumption of full mediation can be tested since the growth predictor model is nested within the direct effect model. The difference between both ways of modeling time-invariant covariates is illustrated by Figs. 5 and 6, which present a simplified presentation of a growth curve model from both perspectives.

Before the analysis can be performed, however, a decision needs to be made about which measurement level is regressed on the time-invariant covariate. That is, school investment and self-confidence were measured with multiple indicators, and their growth models were specified as second-order factor models on the first-order common factors per occasion. Thus, in principle, one can regress the first order factors on the time-invariant covariate, as well as the observed indicators.

For language there appeared to be no problem since its growth curve model is not based on a longitudinal common factor model. The growth parameters account for the structure that is present in the observed indicators. For self-confidence and school investment, however, the growth

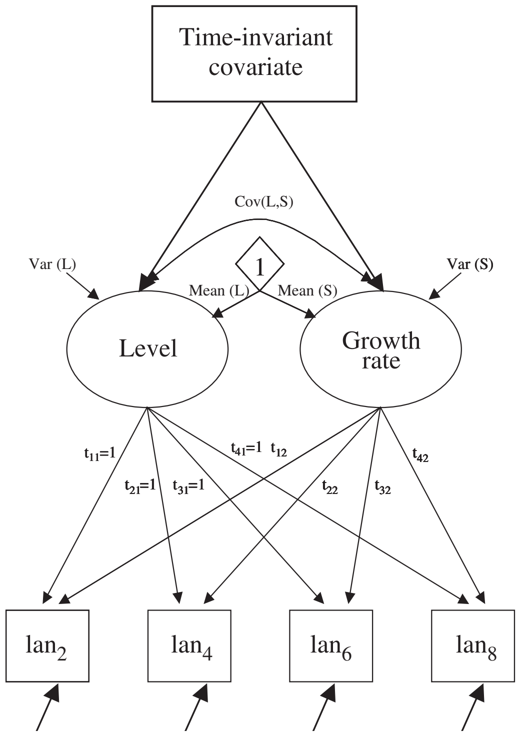

Fig. 5. Growth predictor model.

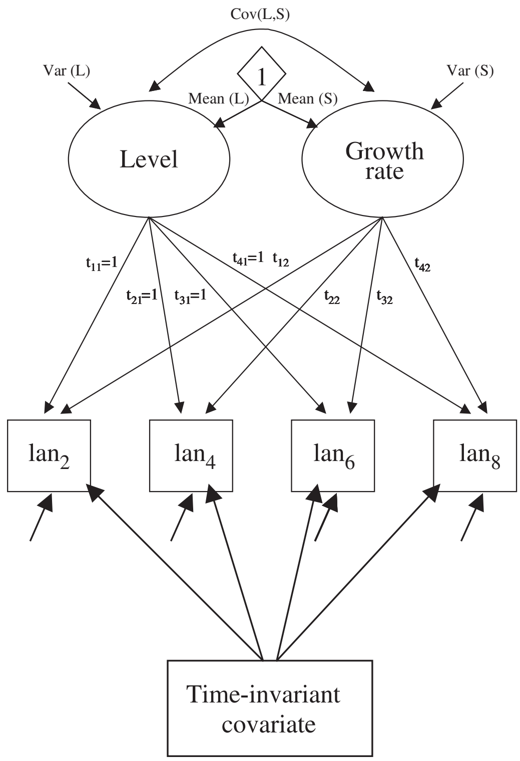

Fig. 6. Direct effect model.

parameters account for the structure of the first order factors. If just the first order factors were regressed on intelligence, it could happen that the residuals of the observed indicators still contain variance that could be, but is not accounted for by intelligence. Therefore, it was decided to regress the observed indicators of all three processes on intelligence. Doing this, the largest effect of the intelligence is expected; at least larger than if the first order factors were regressed on intelligence.

The direct effect multivariate growth curve model, with intelligence as the time-invariant covariate provided the following fit measures: $\chi^2(166)=270.31$, $p=.00$, RMSEA=.014. When this model is contrasted to the growth predictor model ($\chi^2(180)=360.49$, $p=.00$, RMSEA=.018), a chi-square difference of $\Delta\chi^2(14)=90.18$ for the test of the assumption of full mediation is obtained. In other words, the growth predictor model has to be rejected for these data, because of a violation of the assumption of full mediation. Intelligence appears to be related not only to the growth parameters, but also to variance of the observed indicators that is not accounted for by the growth parameters. With the exception of the second measure of self-confidence at the first measurement occasion, all direct effects were significantly different from zero, ranging in magnitude from .05 to .48 after standardization.

The additional analysis presented here points out that the model on which the final conclusions of Stoel, Peetsma and Roeleveld (2003) should have been a direct effect model. To investigate whether this had consequences for the relevant parameter estimates, the correlations between the

Table 4
Comparison of the correlations, and proportions variance explained, between growth parameters of the direct effect model and the growth predictor model

	1.	2.	3.	4.	5.	6.
Proportion variance explained:						
Direct effect model	.35	.36	.00	.25	.24	.00
Growth predictor model	.29	.08	.00	.06	.23	.00
1. Level (language)		0	.14**	0	.24**	0
2. Growth (language)	0		−.16**	.36*	.06	.29**
3. Level (self-confidence)	.12**	−.15**		−.44**	0	0
4. Growth (self-confidence)	0	.35**	−.42**		.16**	0
5. Level (school investment)	.18**	.06	0	.14**		0
6. Growth (school investment)	0	.27**	0	0	0	

**$p < .01$. The sub-diagonal elements present the correlations of the growth parameters of the direct effect model, and the upper diagonal elements correspond to the growth predictor model.

growth parameters of the growth predictor model are now compared to their counterparts of the direct effect model. Since the direct effect model did not provide estimates of the effect of intelligence on each of the growth parameters, the proportions of variance explained of the growth parameters were computed by hand, and compared to the proportions of variance explained in the growth predictor model. Table 4 presents the relevant information. The first rows present the proportion of variance explained, the sub-diagonal elements present the correlations of the growth parameters of the direct effect model, and the upper-diagonal elements correspond to the growth predictor model.

Table 4 shows that the structure of the correlations between the growth parameters remains the same. Although the magnitude of the correlations is slightly changed, these changes are not of such importance that they would have led to different conclusion. What does change, however, are the proportions variance explained by the growth parameters by intelligence. This change is most pronounced for the growth rates of language acquisition and self-confidence for which the proportion variance explained change from, respectively .08 and .06, to .36 and .25. Although it seems that intelligence now has a larger effect on the development of self-confidence and language acquisition, the correlations between the growth parameters of self-confidence and school investment are still significant. So this model also leads to the conclusion that there is an effect of the development in the motivational variables on the development of language acquisition after the effect of intelligence has been controlled for.

5. Discussion and conclusion

The secondary analyses presented in this paper provide a nice illustration of the implications of the conclusions of Stoel, Van den Wittenboer and Hox (2004a,b) for an empirical data set. They showed that standard modeling practices may sometimes lead to incorrect conclusion regarding the concepts under investigation, and that ideally alternative modeling possibilities should be considered in order to check the adequacy of the standard practice. The aim of this paper was to apply the results of Stoel, Van den Wittenboer and Hox to the analyses of Stoel, Peetsma and Roeleveld (2003).

The first issue we focused on was related to the measurement invariance of the repeatedly measured concept of school investment and motivation. Stoel, Van den Wittenboer and Hox (2004b) found that changing the reference indicator in a higher-order latent growth curve model, under the condition of partial measurement invariance, may lead to a change in parameter estimates and model fit, and thus possibly to a change in the substantive conclusions. This finding may be troublesome since reference indicators are often chosen arbitrarily. By changing the reference indicators as they were used by Stoel, Peetsma and Roeleveld (2003), the current analysis showed that in this particular situation the impact of a change in the reference indicator under partial measurement invariance was not very important. Though model fit improved slightly, the parameter estimates and substantive conclusions remained the same.

Regarding the section on modeling the time-invariant covariate we would like to add that if the assumption of full mediation is violated, the effect of a time-invariant covariate on the latent growth parameters will be biased. This argues against the standard practice in latent growth curve modeling, as well as longitudinal multilevel regression analysis, to model the time-invariant covariate with direct effects on the growth parameters without testing the assumption of full mediation. Again, in this specific case the results of the direct effect model were not radically different from the results of the growth predictor model that were reported earlier, except for the pleasing fact that the model fit of the direct effect model was much better. However, as was shown in Stoel, Van den Wittenboer and Hox (2004b), the parameter estimates of the growth predictor model may in other cases be drastically different from their population values, emphasizing the need for an explicit test of the assumption of full mediation.

In addition, it is interesting to note that if there was a significant effect on a growth parameter in the growth predictor model, the effect was underestimated. However, this cannot be regarded as a rule, since the analyses of Stoel, Van den Wittenboer and Hox (2004a) have shown that the effect on the growth parameters may also be overestimated in certain instances. Whether or not the effect of a time-invariant covariate is overestimated, or underestimated, depends on the violation of the assumption of full mediation. Furthermore, if the assumption is violated, it is presumably the relative strength of the true direct effects (i.e. the relation between the covariate and the time specific residuals) that determines the bias. In our case this is the relation between intelligence and the observed indicators of language acquisition, school investment and self-confidence. A simulation study might shed further light on the specific conditions for over and underestimation of the effect of a time-invariant covariate.

The question rises if the argument can be extended to the common factor model in general. That is, shouldn't the assumption of full mediation always be tested before a covariate is modeled as a predictor of latent variables? Or even stronger, if any effect between latent variables is modeled, shouldn't one firstly regress the indicators of the outcome on the predictor? The answer is yet hard to give, since things are a bit different here. Contrary to a (linear) latent growth curve model, factor loadings are in general not constrained to known values. It were exactly these constrained factor loadings that facilitated the proof that the growth predictor model is nested within the direct effect model. These issues, concerning the generalization to the common factor model in general will be the topic of future work.

Appendix A. Estimated covariance matrix and means vector

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.
1. self21	.714																					
2. self22	.418	.785																				
3. self41	.150	.126	.584																			
4. self42	.122	.152	.372	.722																		
5. self61	.094	.081	.117	.115	.601																	
6. self62	.094	.116	.099	.161	.420	.777																
7. self81	.095	.095	.120	.162	.121	.186	.734															
8. self82	.071	.110	.071	.178	.138	.254	.505	.881														
9. inv21	.079	.125	.003	.020	.028	.036	.063	.074	.926													
10. inv22	.029	.090	.027	.029	.038	.056	.081	.069	.578	.851												
11. inv431	.014	.030	.077	.088	.037	.059	.032	.054	.240	.231	.920											
12. inv42	.011	.023	.067	.080	.074	.093	.078	.095	.281	.278	.597	1.075										
13. inv61	-.044	-.027	-.022	-.006	.086	.116	.035	.052	.237	.274	.322	.367	1.067									
14. inv62	-.029	-.016	.027	.030	.118	.129	.086	.078	.261	.300	.364	.467	.664	1.088								
15. inv81	-.025	-.035	-.024	.017	.039	.065	.141	.140	.205	.222	.279	.345	.383	.400	1.104							
16. inv82	-.035	-.008	.031	.061	.077	.092	.155	.156	.228	.269	.292	.412	.379	.491	.692	1.078						
17. lan2	.077	.055	.090	.082	.078	.074	.051	.022	.170	.160	.150	.173	.148	.140	.182	.148	1.22					
18. lan4	.042	.010	.055	.041	.068	.090	.109	.096	.161	.195	.178	.219	.211	.210	.197	.174	.563	1.511				
19. lan6	.053	.075	.064	.064	.109	.124	.149	.120	.224	.209	.210	.269	.265	.284	.275	.310	.553	.757	1.405			
20. lan8	.016	.037	.063	.035	.113	.125	.190	.150	.212	.228	.212	.253	.257	.293	.307	.316	.537	.695	.853	1.320		
21. Figure	.108	.113	.087	.073	.230	.236	.320	.247	.446	.412	.298	.431	.384	.501	.443	.506	.539	.722	.924	.865	7.518	
22. Excl.	.125	.058	.118	.133	.174	.206	.296	.265	.425	.489	.327	.474	.568	.568	.486	.557	.747	.862	1.027	.920	2.984	6.910
Means	3.87	3.58	3.81	3.59	3.83	3.61	3.82	3.59	.358	3.62	3.53	3.12	3.44	3.21	3.45	3.25	32.43	34.71	36.02	37.25	14.43	11.67

Care should be taken in replication of the current analyses by means of these means and covariances, since they represent the estimated moments provided by M*plus* as a by-product of the FIML estimation method. FIML estimation requires raw data, which are available upon request from the first author. N_{total}=2693 for Language Ability (lan); N_{total}=2660 for Self-confidence (self) and School investment (inv). The difference in total sample size is due the fact that a subsample of n=33 subjects have no observations for Self-confidence and School investment.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two step approach. *Psychological Bulletin*, *103*, 411–423.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, *21*, 230–258.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and meanstructures: the issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456–466.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Garst, H., Frese, M., & Molenaar, P. C. M. (2000). The temporal factor of change in stressor–strain relationships: A growth curve model on a longitudinal study in East Germany. *Journal of Applied Psychology*, *85*, 417–438.

Hancock, G. R., Kuo, W. L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, *8*, 470–489.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent growth curve models. *Multivariate Behavioral Research*, *32*, 215–253.

McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics*, *16*(1), 163–200.

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

McArdle, J. J. (1989). Structural modeling experiments using multiple growth functions. In P. Ackerman, R. Kanfer, & R. Cudeck (Eds.), *Learning and individual differences: Abilities, Motivation and Methodology* (pp. 71–117). Hillsdale, NJ: Lawrence Erlbaum Associates.

Meredith, W. M., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*, 431–462.

Muthén, B., & Khoo, S. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, *10*, 73–102.

Muthén, L.K. and Muthén, B.O. (1998). M*plus* 1.04 [Computer software]. Los Angeles: Muthén and Muthén.

Oort, F. J. (2001). Three-mode models for multivariate longitudinal data. *British Journal of Mathematical and Statistical Psychology*, *54*, 49–78.

Pentz, M. A., & Chou, Ch. P. (1994). Measurement invariance in longitudinal clinical research assuming change from development and intervention. *Journal of Consulting and Clinical Psychology*, *62*, 450–462.

Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins, & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179–200).

Stoel, R. D., Peetsma, T. T. D., & Roeleveld, J. (2003). Relations between the development of school investment, self-confidence and language achievement in elementary education: A multivariate latent growth curve approach. *Learning and Individual Differences*, *13*, 313–333.

Stoel, R. D., Van den Wittenboer, G., & Hox, J. J. (2004a). Including time-invariant covariates in the latent growth curve model. *Structural Equation Modeling*, *11*, 155–167.

Stoel, R. D., Van den Wittenboer, G., & Hox, J. J. (2004b). Methodological issues in the application of the latent growth curve model. In K. van Montfort, H. Oud, & A. Satorra (Eds.), *Recent developments on structural equation modeling: Theory and applications* (pp. 241–262). Amsterdam: Kluwer Academic Press.

Verhelst, N. D., Glas, C. A. W., & Verstralen, H.H.F.M. (1993). OPLM: One parameter logistic model. *Computer program and manual*. Arnhem: CITO.

Vierke, H. (1995). *De PRIMA toetsen gecalibreerd [Calibration of the PRIMA tests]*. University of Nijmegen, Nijmegen: ITS.

Vonesh, E. F., & Chinchilli, V. M. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*, 363–381.