academicJournals

Vol. 11(15), pp. 1440-1448, 10 August, 2016
DOI: 10.5897/ERR2016.2773
Article Number: FE0F30459869
ISSN 1990-3839

http://www.academicjournals.org/ERR

Educational Research and Reviews

Full Length Research Paper

Measurement equivalence of teachers' sense of efficacy scale using latent growth methods

T. Oğuz Başokçu* and T. Öğretmen

Department of Educational Sciences, Division of Measurement and Evaluation in Education, Faculty of Education, Ege University , İzmir, Turkey.

Received 19 March, 2016; Accepted 14 June, 2016

This study is based on the application of latent growth modeling, which is one of structural equation models on real data. Teachers' Sense of Efficacy Scale (TSES), which was previously adapted into Turkish was administered to 200 preservice teachers at different time intervals for three times and study data was collected. Measurement equivalence of TSES was analyzed over this data using second-order latent growth models. For this aim, whether the scale achieved measurement equivalence at metric and scalar level was tested. The findings support that Teachers' Efficacy Scale has measurement equivalence in time. Time intervals between the applications of the scale were associated with school experience (internship programs) of preservice teachers. The applications were carried out before, during and after internship. It was found that preservice teachers' sense of efficacy increased in time and that as school experience increased belief towards professional sense of efficacy developed.

Key words: Latent growth models, measurement equivalence, longitudinal study, teachers' sense of efficacy.

INTRODUCTION

Latent growth models (LGM) among structural equation models, is used to analyze time-dependent change. These models were first introduced by Rao (1958) and Tucker (1958) as the basis of longitudinal factor analysis and were later developed and corrected by Meredith and Tisak (1984, 1990), McArdle (1988), McArdle and Epstein (1987) and Muthen (1991). Longitudinal analyses use different statistical techniques for data analysis. One of these methods is the approach, which deals with "raw change" in scores. Analysis of variance (ANOVA) and multiple regression methods analyze the difference between the first and second time measurements as a function of individual or group characteristic (Curran and Muthen, 1999). Another alternative approach involves the analysis of "residual variance". Analysis of covariance (ANCOVA) is one of the typical applications of this method. ANOVA only analyzes "raw change;" in other words, the change in mean and finally change in each individual is indicated as error variance. However, this "error variance" in fact, gives important information about the quality of change. For this reason, researchers studied methods to better explain the nature of change and differentiation of the individual (Duncan and Duncan 2004). Time series analysis methods model change in data, rather than the change in structure. Although LGM, which are developed as a result of these studies, is a

*Corresponding author. E-mail: tahsin.oguz.basokcu@ege.edu.tr.

stronger method in statistical terms, in studies which are based on the comparison of two methods, it gives a higher fit between the underlying theory of the analysis method and the statistical method (Preacher et al., 2008).

Latent growth model

LGM offer a more suitable methodology for longitudinal studies, which measure the change in the latent structure consisting of manifest variables. The variables observed in these types of studies are generally obtained from total or mean scores obtained in time of an item set under a single structure. The obtained observed variables are used in identification of first order LGM. In first-order LGM, the model fits to the vector and covariance matrix of mean observed variables scores which were reformed at each time for each individual (Sayer and Cumsille, 2001; Andruff et al., 2009; Sessoms and Finney, 2015; Kim and Willson, 2014).

While calculating time-dependent variance in LGM, two parameters of latent growth factor were determined. These parameters are first situation and change ratio. In model function, the first situation corresponds to intercept and variance ratio corresponds to slope (Fraine et al., 2007).

Considering that the markers observed for first-order LGM are *X, W*…and *Y* and the sum of these items point out to *V* observed variable, *V=X+W+….+Y*.

For first-order LGM, it is expressed as follows;

$$V_{(t)n} = f_{0n} + \beta_{(t)} f_{sn} + e_{(t)n},$$

Where *V* indicates observed variable score observed at *t* time for *n* people while $f_{0n}$and f_{sn} show intercept and slope values. Apart from this, β shows the basic coefficient determined from the shape of the slope and *e* indicate residual terms for *n* people at *t* time which cannot be explained with intercept and slope values (Ferrer et al., 2008).

Since first-order LGM is easily calculated and interpreted, they are commonly used in a lot of fields. On the other hand, easy adaptation of model is one of the strong aspects of first-order LGM. However, the principle limitation of this model is that the same latent structure is evaluated for the measurements at each time level. Presence of a single indicator for each measurement time makes it impossible to apply standard procedure to identify factorial invariance. For this reason, researchers cannot calculate but only estimate the variance, which is the effect of the observed variable on the latent invariance dimension for each measurement time (Ferrer et al., 2008).

Second-order LGM is an approach which allows for modeling of the change of latent level in time. This model is a multiple variable extension of first degree LGM. Its most important advantage is that factorial invariance of model parameters can be calculated for the measurements at each time level. When factorial invariance is kept constant, it can inform the researcher that the same latent dimension is measured as the same at each time level. If factorial invariance is achieved, in that case, intercept and slope parameters of the model can be calculated. For example, a model with three variables can be written as follows;

$$X_{(t)n} = \tau_x + \lambda_x f_{(x)n} + e_{x(t)n},$$
$$W_{(t)n} = \tau_w + \lambda_w f_{(x)n} + e_{w(t)n}, ve$$
$$Y_{(t)n} = \tau_y + \lambda_y f_{(x)n} + e_{y(t)n},$$
$$f_{(t)n} = f_{0n} + \beta_{(t)} f_{sn} + z_{(t)n}\ ,$$

In the equation above, observed variables of *n* people at *t* time are indicated with X, W and Y. In the equation, τ indicates the intercept, λ indicates factor load or the indicator of linear slope for factor *f*, *e* indicates single factor score, f_{0n} and f_s indicate intercept and slope values of *f* factor for *n* people and β indicate basis coefficients determined from the shape of the slope and $z_{(t)}$expresses distribution term of*f* factor at *t* time. Diagram of this model is presented in Figure 1. In the figure, observed variables are indicated with square, while latent variables are indicated with circle. On the other hand, triangles represent the intercept with which means and slopes are estimated.

Three prerequisites should be satisfied to make LGM analysis (Kline, 2005). First of all, measurement at constant (at equal intervals) level at minimum three different situations belonging to one independent variable is required. Secondly, the measurements should be collected from the individuals simultaneously and finally measurement equivalence should be achieved (Dural et al., 2011). Achieving measurement equivalence means that the measurements observed at different time points measured the same structure.

Measurement equivalence for LGM

Longitudinal measurement equivalence or time-dependent measurement equivalence indicate the status of measurement values obtained from the same measurement tool at different times. Achieving measurement equivalence means that the structure formed by the observed variables by latent structure measured at different times is the same at each measurement time (Drasgow, 1987; Meredith, 1993; Hancock et al., 2001; Ferrer et al., 2008).

Factorial equivalence whose scope was determined by Meredith (1964, 1993) was analyzed by taking measurement equivalence of common factor model as a reference in general terms (Stoel et al., 2004, Ferrer et al., 2008). In previous studies, factorial equivalence was divided into two categories, which are metric equivalence and non-metric equivalence (Horn et al., 1983; Meredith, 1993; Widaman and Reise, 1997). Metric equivalence is hierarchically analyzed in three steps which are weak,

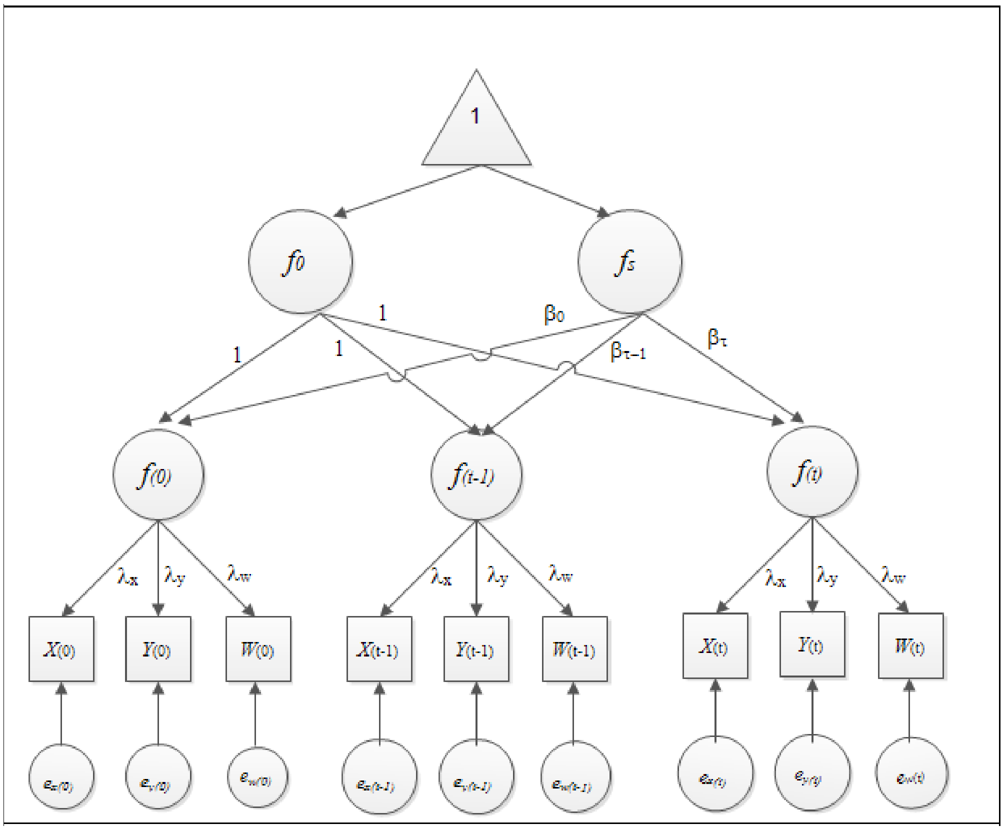

Figure 1. Second order LGM.

strong and strict factorial equivalence (Ferrer et al., 2008).

Non-metric equivalence means similar properties of indicators of latent structure without numerical data for parameter estimates. Weak factorial equivalence points out those factor loads of each observed variable take time-dependent values. On the other hand, strong factorial equivalence refers to adding invariance of intercept value of each indicator to weak equivalence. In strict factorial equivalence, in addition to factor loads and intercept value, it is expressed that specific variance of each indicator also changes in time (Ferrer et al., 2008). Analysis of measurement equivalence using second-order latent growth models only test weak and strong equivalences; after achieving strong equivalence, strict equivalence analysis cannot be generally performed. In other words, achieving scalar (strong) invariance is considered as adequate for measurement equivalence (Vandenberg and Lance, 2000; Wu et al., 2007).

Chi-square "χ^2" values and Comparative Fit Index (CFI) determined from two models are used to determine whether measurement equivalence is achieved or not (Byrne and Stewart, 2006; Vandenberg and Lance, 2000; Wu et al., 2007). $\Delta\chi^2$ and ΔCFI are calculated taking the difference of χ^2 and CFI values belonging to two models. While testing statistical significance of the obtained $\Delta\chi^2$, this value is compared to critical chi-square value corresponding to the difference between degrees of freedom of two models. Obtaining a statistically insignificant $\Delta\chi^2$ value as a result of this comparison indicates that measurement equivalence is achieved. Although no statistical significance test can be made for ΔCFI, if the ΔCFI value obtained from the comparison of two models is equal to or smaller than -0.01 value can be used as a proof of achieved measurement equivalence (Byrne and Stewart, 2006; Vandenberg and Lance, 2000; Wu et al., 2007).

Aim

This study examined time-dependent measurement equivalence of second-order LGM using real data. The teachers' sense of efficacy scale's weak and strong measurement equivalence in second order LGM was tested. Besides, the measurement tool's longitudinal measurement equivalence was examined. This study also investigated further evidence for the construct validity of the three factor subscale scores via analyzing the measurement equivalence of TSES using second-order latent growth models. On the other hand, we tried to analyze how Teachers' Sense of Efficacy changed as the teachers' experience increased.

METHODS

Type of the study

Since the study is based on the examination of data on the practical use of LGM models, it is a descriptive study. However, as the measurements of Teachers' Sense of Efficacy were taken at different time intervals, it can be considered as longitudinal study.

Study group

Sampling of the study included preservice teachers enrolled in Ege University, Program of Teaching (Pedagogic Training for Preservice Teachers). Real data was obtained from repetitive measurement of Teachers' Sense of Efficacy Scale on 200 students who were randomly selected from a total of 400 students. The sampling consisted of a total of 200 students with bachelor's degree, of whom 128 (64%) were female and 72 (36%) were male.

Measurement tool

Original form of Teachers' Sense of Efficacy Scale-TSES was developed in "Self Efficacy in Learning and Teaching" seminar held in Ohio State University. New items were added based on self-efficacy scale of Bandura and a 24-item scale consisting of 9-item Likert scale was designed (Tschannen-Moran and Hoy, 2001). The reliability for the 24-item scale was 0.94 and construct validity evidence obtained by authors.

Original scale was adapted in Turkish in 2005 after completing validity and reliability studies (Çapa et al., 2005). The Turkish adaptation of the scale contained 3 sub-dimensions and 24 items like in the original one. Sub-dimensions of the scale measure teacher competencies in "using educational strategies", "classroom management" and "providing student participation".

Teachers' sense of efficacy refers to belief or perception of a teacher that he/she can produce desired learning products even in case of the presence of generally problematic students with low motivation (Tschannen-Moran and Woolfok-Hoy, 2001). The characteristics of individuals with sense of efficacy include struggling with problems and having a determined attitude towards problem-solving, in other words, loyalty to new targets (Bandura, 1997; Scholz et al., 2002).

Cronbach's alpha coefficient calculated for Turkish adaptation of Teachers' Sense of Efficacy was found to be .93. However, it was calculated as .84 for "using educational strategies"; .84 for "classroom management" and 0.86 for "providing students participation" (Capa et al., 2005).

Total scores obtained from Teachers' Sense of Efficacy Scale gives information about professional sense of efficacy of teachers or preservice teachers. High scores received from the scale indicate higher sense of efficacy for teaching profession while low scores indicate lower sense of efficacy. It also provides information about their levels in classroom management, providing students participation and using teaching strategies which are sub-dimensions of teachers' sense of efficacy.

Procedure

Teachers' Sense of Efficacy Scale was administered to a total of 200 preservice teachers who were selected as unbiased in three replications before (the first application), during (in the middle of internship) and after the internship (the third application) during 2012/2013 spring semester. This application aimed to examine how teachers' sense of efficacy varied before having any experience; after having a certain degree of school experience and after completing internship programs. The data were gathered at 60-day intervals.

Analyses

M-Plus 5.1 (Muthen and Muthen, 2008) software was used for the analysis of latent growth models. In the present study, measurement equivalence in second degree LGM was analyzed using basic LGM, weak measurement equivalence in LGM and strong measurement equivalence in LGM.

RESULTS

Data analyses

Basic LGM model in Table 1 indicates a situation with no limitation to test measurement equivalence. Indices obtained from basic LGM in the analysis indicate a good fit between the model and data. In other words, measurements of latent variable at three time points show a linear increase. However, for the situation which achieved measurement equivalence, the difference between latent variable should be analyzed. Weak and strong equivalents should be tested. Table 1 presents the results of basic model, weak invariance model and strong invariance model.

Basic LGM model in Table 1 indicates a situation with no limitation to test measurement equivalence. Indices obtained from basic LGM in the analysis indicate a good fit between the model and data. In other words, measurements of latent variable at three time points show a linear increase. However, for the situation which achieved measurement equivalence, the difference between latent variable should be analyzed. Weak and strong equivalents should be tested.

Factor loadings were restricted in weak equivalence degree included in analyses. For this model, measurement equivalent was determined by subtracting χ^2 and CFI values obtained for the first model from χ^2 and CFI values. Analysis of $\Delta\chi^2$ and ΔCFI in Table 1 shows that there was no significant deterioration in the model. This means that Teachers' Sense of Efficacy Scale had weak (metric) measurement equivalence in LGM. In addition to the limitations in strong equivalence weak equivalence analyses for the scale, analysis was made by restricting intercept values. As indicated in Table 1, ΔCFI and $\Delta\chi^2$ values obtained from the difference of CFI and χ^2 values of weak and strong invariance models indicate that there is no deterioration in the model. This situation proves that condition of scalar (strong) equivalence was achieved in second level LGM for Teachers' Sense of Efficacy Scale.

The Appendix 1 includes parameter estimates and reports results of factor loads. As indicated in program

Table 1. Goodness of fit and model difference statistics of measurement equivalence.

Parameter	χ^2	sd	SRMR	CFI	$\Delta\chi^2$	Δsd*	ΔCFI
Basic LGM	118.915	25	0.051	0.915	–	–	–
Weak equivalence in LGM	120.857	29	0.069	0.917	1.942	4	0.002
Strong equivalence in LGM	122.365	33	0.070	0.919	1.508	4	0.002

Δsd* = 4 critical value; χ^2 = 13.28 at p = 0.01 level.

output, since first indicators off latent variable are fixed to 1 in the model as reference variables, no parameter estimation can be made. On the other hand, for two other indicators of latent variable, parameter estimations are given by restricting factor loads in such a way to be equivalent at three time points. "S WITH I" section in the Appendix 1 show covariance value for latent growth factors and it is understood that this coefficient is -19.06 (p = 0.00). In standardized results, correlation coefficient was -0.89 among latent growth factors which corresponds to this value. The section "Means" in parameter estimates indicates means of latent growth factors and here it is understood that mean first situation latent growth factor was estimated as zero. Main reason for this is that in second degree LGM, constant values of observed variables and mean of the first situation latent growth factor were not estimated simultaneously. Estimated mean for the change ratio of latent growth factor is 0.27 (p = 0.00). The section "Intercepts" which reports constant values for observed variables include restricted parameter estimations with equivalent constant values at three time points. Estimated variances of the first situation and change ratio factors (under the title "variances") were found to be 30.62 (p = 0.00) and 14.70 (p = 0.00) respectively. The last chapter of the Appendix 1 reports error variance values for observed (y11-y33) and latent (f1-f3) variables under the title of "Residual Variances". R^2 values which represent explanation ratio of the model of observed and latent variables were calculated as 0.71, 0.75, 0.70, 0.62, 0.68, 0.65, .81, 0.77 and 0.70 for observed variables and 0.33, 0.72 and 0.51 for latent variables, respectively.

DISCUSSION

In the present study, linear change of teachers' efficacy in time was analyzed by using real application data set within the scope of second order LGM; measurement equivalent was tested as metric and scalar level and findings of the analyses were presented. Furthermore, model parameters in the outputs were explained.

The fact that estimated mean of latent growth factor of variance ratio was statistically significant points out to a time-dependent linear change in terms of the handled characteristics. The fact that mean of latent growth factor of variance ratio was statistically significant indicate that teachers' sense of efficacy showed a time-dependent variance. Since the obtained mean value is positive, the mentioned variance can be interpreted as preservice teachers" increased teachers' sense of efficacy in time.

Statistically significant result obtained for the variance of the first situation latent growth factor means that the individuals in the sampling was not a heterogeneous group in terms of teachers' sense of efficacy levels. On the other hand, statistically significant variance of latent growth factor of change ratio points out that the individuals in the sampling different from each other in time, in terms of teachers' sense of efficacy levels.

The relationship between latent growth factors (in other words obtaining a statistically significant covariance value) indicates that the increase the individuals with high levels of teachers' sense of efficacy at the beginning show in time was higher than that of the individuals with low levels at the beginning (Bollen and Curran, 2006; Welch, 2007). Despite this, since there is a negative relationship, it indicates that teachers "sense of efficacy levels of the individuals with lower beginning level increased more than those who had higher teachers" sense of efficacy levels at the beginning. Finally, R^2 values reported in standardized results for the estimated error variances of the observed (y11-y33) and latent (f1-f3) variable in the model are used to interpret the extent the variance in observed and latent variables are explained by the model. According to this, variance model explained 71% of variance in observed variables and 52% of the variance in latent variables on average.

Conclusion

Various research problems about sensory characteristics of individuals require analysis of behavior in time. The results of this study indicated that Teachers' Sense of Efficacy Scale had weak (metric) measurement equivalence and scalar (strong) equivalence in LGM. Considering that teachers' sense of efficacy can change as the experience of preservice teacher increases, widespread use of this type of research particularly in the field of education will significantly contribute to our knowledge of teacher training. Teachers' sense of efficacy form was administered to the students enrolled in program of preservice training, from whom data was collected, at three different times: Before the start of teaching application (the first application); in the middle of semester (the second application) and after the completion of application training (the third application). This made it possible to investigate the change in professional sense of efficacy of the students in time depending on experience. It was found that as the school experience of the students with low sense of efficacy at the beginning, professional sense of efficacy levels increased. As implied by Mulholland and Wallace in 2001, the findings reveal that as their school experience increased, the students with high sense of efficacy score at the beginning shows lower development than those who had low sense of efficacy. Based on our findings it can be stated that measuring sense of efficacy in individuals with no school experience can be misleading. In other words, it can be suggested that professional sense of efficacy becomes more stable with experience.

A number of research should be conducted in future studies: First, further research on psychometric properties of the Teachers' Sense of Efficacy Scale needs to be analyzed. Second, the scale quality should be tested for investigation of the relationships between in-service teachers across different settings and different subject-areas. Finally, the relationships between teacher characteristics and teachers' efficacy judgments should be analyzed.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

Andruff H, Carraro N, Thompson A, Gaudreau P, Louvet B (2009). Latent class growth modelling: A tutorial.Tutorials in Quantitative Methods for Psychol. 5(1):11-24.

Bandura A (1997). Self-efficacy: The exercise of control. New York: Freeman.

Bollen KA, Curran PJ (2006). Latent curve models: A structural equation perspective. New Jersey: A John Wiley & Sons, Inc., Publication.

Byrne BM, Stewart SM (2006). The MACS approach to testing for multigroup invariance of a second – order structure: A walk through the process. Structural Equating Modeling, 13(2):287-321.

Capa Y, Cakiroglu J, Sarikaya H (2005). The development and validation of a Turkish version of teachers' sense of efficacy scale. Egitim ve Bilim (Education and Science), 30(137):74-81.

Curran PJ, Muthén BO (1999). The application of latent curve analysis to testing developmental theories in intervention research. Am. J. Community Psychol. 27:567-595.

Drasgow F (1987). Study of the measurement bias of two standardized psychological tests. J. Appl. Psychol. 72:19-29.

Duncan TE, Duncan SC (2004). An introduction to latent growth curve modeling. Behavior Therapy, 35(2):333-363.

Dural S, Somer O, Korkmaz M, Can S, Öğretmen T (2011). İkinci Drerece Örtük Gelişme Modelleri ve Ölçme Eşdeğerliği. Eğitim ve Bilim, 36(161):51-63.

Fraine BD, Damme JV, Onghena P (2007). A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. Contemporary Educ. Psychol. 32(1):132-150. doi: http://dx.doi.org/10.1016/j.cedpsych.2006.10.005

Ferrer E, Balluerka N, Widaman KF (2008). Factorial invariance and the specification of the second-order latent growth models. Methodology 4:22-36.

Hancock GR, Kuo W, Lawrence FR. (2001). An illustration of second-order latent growth models. Structural Equation Modeling 8(3):470-489.

Horn JL, McArdle JJ, Mason R. (1983). When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. Southern Psychologist 1:179-188.

Kim ES, Willson VL (2014). Measurement Invariance Across Groups in Latent Growth Modeling, Structural Equation Modeling: A Multidisciplinary J. 21(3):408-424 DOI: 10.1080/10705511.2014.915374

Kline RB (2005). Principles and practice of structural equation modeling. New York: The Guilford Press.

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

McArdle JJ, Epstein D (1987). Latent growth curves within developmental structural equation models. Child Dev. 58:110-133.

Meredith WM (1964). Notes on factorial invariance. Psychometrika 29:177-185.

Meredith WM (1993). Measurement invariance, factor analysis, and factorial invariance. Psychometrika 58:525-543.

Meredith W, Tisak J (1984). "Tuckerizing" curves. Paper presented at the annual meeting of the Psychometric Society July, Santa Barbara, CA.

Meredith W, Tisak J (1990). Latent curve analysis. Psychometrika, 55:107-122.

Mulholland J, Wallace J (2001). Teacher induction and elementary science teaching: enhancing self-efficacy. Teach. Teacher Educ. 17:243-261.

Muthén B (1991). Analysis of longitudinal data using latent variable models with varying parameters. In: L. M. Collins & J. L. Horn (Eds.), Best methods for analysis of change. Washington DC: Am. Psychol. Assoc. pp. 1-17.

Preacher K, Wichman A, MacCallum R, Briggs N (2008). Latent Growth Curve Modeling, Series: Quantitative Applications in the Social Sciences, SAGE Publications, California

Rao CR (1958). Some statistical methods for the comparison of growth curves. Biometrics 14:1-17.

Sayer AG, Cumsille PE (2001). Second-order latent growth models. In: Collins, LM.; Sayer, AG., editors. New methods for the analysis of change. Washington, D. C.: Am. Psychol. Assoc. pp. 179-200.

Sessoms J, Finney Sara J (2015). Measuring and modeling change in examinee effort on low-stakes tests across testing occasions. Int. J. Testing 15(4):356-388, DOI: 10.1080/15305058.2015.1034866

Scholz U, Gutierrez-Dona B, Sud S, Schwarzer R (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. Euro. J. Psychol. Assessment 18(3):242-251.

Stoel RD, van den Wittenboer G, Hox JJ (2004). Methodological issues in the application of the latent growth curve model. In: van Montfort, K.; Oud, H.; Satorra, A., editors. Recent developments on structural equation modeling: Theory and applications. Amsterdam: Kluwer Academic Press pp. 241-262.

Tschannen-Moran M, Woolfolk Hoy A (2001). Teacher efficacy: Capturing an elusive construct. Teacher Educ. 17(7):783-805.

Tucker LR (1958). Determination of parameters of a functional relation by factor analysis. Psychometrika 23:19-23.

Vandenberg RJ, Lance CE (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods 3:4-69.

Welch GW (2007). Model fit and interpretation of non-linear latent growth curve models. Unpublished dissertation thesis, University of Pittsburg.

Widaman KF, Reise SP (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In: Bryant, KJ. Windle, M.; West, SG., editors. The science of prevention: Methodological advances from alcohol and substance abuse research. Washington, DC: American Psychological Association; pp. 281-324.

Wu AD, Li Z, Zumbo BD (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. Practical Assessment, Res. Eval. 12:1-26.

Appendix 1. Second order LGM' M-Plus output for standardized solutions of parameter estimations.

Variable				Two-Tailed
	Estimate	S.E.	Est./S.E.	P-Value
		STD Standardization		
I				
F1	1.152	0.080	14.390	0.000
F2	1.094	0.092	11.909	0.000
F3	1.055	0.105	10.018	0.000
S				
F1	0.000	0.000	999.000	999.000
F2	0.760	0.069	11.056	0.000
F3	1.466	0.117	12.515	0.000
F1 by				
Y11	4.805	0.341	14.080	0.000
Y21	5.654	0.386	14.648	0.000
Y31	5.730	0.412	13.903	0.000
F2 by				
Y12	5.059	0.419	12.080	0.000
Y22	5.411	0.418	12.955	0.000
Y32	5.567	0.440	12.656	0.000
F3 by				
Y13	5.245	0.334	15.712	0.000
Y23	5.735	0.379	15.145	0.000
Y33	5.527	0.390	14.187	0.000
S with I	-0.896	0.037	-24.373	0.000
Means				
I	0.000	0.000	999.000	999.000
S	0.273	0.087	3.130	0.002
Intercepts				
Y11	53.019	0.411	128.957	0.000
Y21	54.177	0.469	115.616	0.000
Y31	53.234	0.492	108.268	0.000
Y12	53.019	0.411	128.957	0.000
Y22	54.457	0.575	94.703	0.000
Y32	53.181	0.599	88.788	0.000
Y13	53.019	0.411	128.957	0.000
Y23	53.938	0.559	96.532	0.000
Y33	53.241	0.559	95.322	0.000
F1	0.000	0.000	999.000	999.000
F2	0.000	0.000	999.000	999.000
F3	0.000	0.000	999.000	999.000
Variance				
I	1.000	0.000	999.000	999.000
S	1.000	0.000	999.000	999.000
Residual variance				
Y11	9.377	1.406	6.667	0.000
Y21	10.658	1.767	6.030	0.000
Y31	14.182	2.047	6.928	0.000
Y12	15.481	2.323	6.666	0.000
Y22	14.046	2.276	6.172	0.000
Y32	16.619	2.516	6.606	0.000
Y13	6.414	1.220	5.257	0.000
Y23	9.777	1.581	6.183	0.000

Appendix 1. Cont'd.

Y33	12.818	1.718	7.460	0.000
F1	99.00	999.00	999.00	999.00
F2	0.715	0.048	15.023	0.000
F3	0.508	0.146	3.472	0.001
		R-square		
Observed variable				
Y11	0.711	0.047	15.067	0.000
Y21	0.750	0.045	16.632	0.000
Y31	0.698	0.048	14.694	0.000
Y12	0.623	0.060	10.392	0.000
Y22	0.676	0.056	12.030	0.000
Y32	0.651	0.056	11.523	0.000
Y13	0.811	0.039	20.821	0.000
Y23	0.771	0.041	18.922	0.000
Y33	0.704	0.044	16.031	0.000
Latent variable				
F1	0.328	0.079	4.152	0.000
F2	0.285	0.048	5.981	0.000
F3	0.492	0.146	3.359	0.001