

## Structural modelling of reading achievement for Lithuanian students

Aistė ELIJO, Gediminas MURAUSKAS (VU)

e-mail: aistuciuke@hotmail.lt, gediminas.murauskas@maf.vu.lt

### Introduction

The ability to read is regarded as fundamental to all forms of personal learning and intellectual growth. It is one of the most important abilities students acquire during their first years of schooling, and the foundation for learning across all subjects [5]. Therefore, the International Association for the Evaluation of Educational Achievement (IEA) has started a series of international studies aimed at measuring primary school students' reading achievements and searching for the factors that might be related to those achievements. The first round of Progress in International Reading Literacy Study (PIRLS) took place in 2001 with 35 countries participating in it. The international database was made available in the year 2003, and since then researchers around the world have been investigating the relationships between their students' reading literacy skills and a number of social, educational, cultural, and other factors.

The issue of the impact of various factors to the educational achievement is not a new one. To quote but few, Israel *et al.* [2] conclude that both parents' socioeconomic status and social capital available in the family promote child's educational achievement. They further note that community social capital also helps children excel in school, although it makes smaller contribution to academic performance. Jensen and Seltzer [3] show that individual, family and neighborhood factors all influence further education decisions of young Australian students. Thirunarayanan [4] compares students' achievements in different content areas by school location in the United States and concludes that students in central-city schools in the United States perform statistically "significantly worse" in many subject areas than students in suburban schools.

Measure of factors that seem to have influence on reading or other educational achievement is another question for analysis. How is "a socio-economic status" or "a good educational climate at home" measured? To give answers to such questions, the use of latent variables and building structural equation models is often carried out. One example of the use of PIRLS 2001 data in this way is an article by Hansen *et al.* [1] about the latent variable modelling. Here they analyze dimensionality of the student and home background factors related to reading, including home educational resources (HER), early home literacy activities (EHLA), students' reading self concept (SRSC), students' attitudes towards reading (SATR), and parents' attitudes towards

reading (PATR). This article served as an inspiration for us to try out similar measurement models for Lithuanian data as well as to construct a structural equations model (SEM) for the relationship of those factors with the students' reading achievements. As previous research by Elijo [6] showed statistically significant differences between the achievements of students from urban and rural schools in Eastern European countries including Lithuania, we find it of interest to compare the models based on school locality.

### Data

PIRLS sample of Grade 4 Lithuanian students was included in the current study. The constructs and their indicators are presented in Table 1. 24 variables were involved in measuring the constructs. There were at least 5 indicators for each construct.

Differently from Hansen *et al.* [1], we do not include latent variable of students' reading self concept (SRSC) in the model. The reliability of the observed variables comprising SRSC was low. It also appeared that Lithuanian translation of one of the questions was not equivalent to the international English version, and it seemed to be misunderstood by Lithuanian students. The remaining number of questions related to students' reading self concept was insufficient for producing a latent variable construct.

The reliability of the included indices was moderate, varying from .608 (in case of HER) to .651 (for PATR), and it was similar to other countries.

### Results

The structural analysis was carried out with AMOS 5.0 (an acronym for "Analysis of Moment Structures") – powerful and easy-to-use structural equation modeling (SEM) software.

First, the measurement models done on Swedish data by Hansen *et al.* [1] were tested on Lithuanian data. Then, multiple-group analysis of measurement models for urban and rural schools in Lithuania was carried out. Finally, we analyzed full structural equations model for Lithuanian students. Full structural equations model contains: the measurement model, which investigates how indicators are related to the latent variables and structural model, which investigates how latent variables are related to one another and observed variable (students' reading achievements). Unknown parameters were estimated using Maximum likelihood estimation method (MLE).

The measurement model for the  $p$  observed variables, represented by the vector  $x$ , relate the observed variables (indicators) to the underlying latent variables (factors) and may be expressed as

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda \mathbf{f} + \boldsymbol{\varepsilon}.$$

Here

$x$ :  $p$ -dimensional vector of indicators,

$f$ :  $q$ -dimensional vector of underlying latent factors. Assume  $f$  is random such that  $E(f) = 0$  and  $Var(f) = \Phi$ .

$\varepsilon$ :  $p$ -dimensional vector of random error. Assume  $E(\varepsilon) = 0$  and  $Var(\varepsilon) = \Psi$ .

Table 1. Reading related factors and their indicators

Factors	Questions in Questionnaires	Indicators
<b>HER</b> Home Education Resources	StQ19. About how many books are there in your home?	ASBGBOOK(X1)
	StQ20. Do you have any of these things in your home?	ASBGPS1(X2)
	a. Computer	ASBGPS2(X3)
	b. Study desk/table for own use	ASBGPS3(X4)
	c. Books of your very own	ASBGPS4(X5)
	d. Daily newspaper	ASBHCHBK(X6)
	HQ13. About how many children's books are there in your home?	ASBHEDUP(X7)
	HQ14. What is the highest level of education completed by the child's father and mother?	
<b>SATR</b> Student's Attitudes Towards Reading	StQ12. What do you think about reading? Tell how much you agree with each of these statements:	ASBGRST1(X8)
	a. I read only if I have to	ASBGRST2(X9)
	b. I like talking about books with other people	ASBGRST3(X10)
	c. I would be happy if someone gave me a book as a present	ASBGRST4(X11)
	d. I think reading is boring	ASBGRST5(X12)
	e. I need to read very well for the future	ASBGRST6(X13)
	f. I enjoy reading	
<b>EHLA</b> Early Home Literacy Activities	HQ2. Before your child began primary school, how often did you or someone else in your home do the following activities with him or her?	ASBHAC1(X18)
	a. Read books	ASBHAC2(X19)
	b. Tell stories	ASBHAC3(X20)
	c. Sing songs	ASBHAC4(X21)
	d. Play with alphabet toys	ASBHAC6(X22)
	f. Play word games	ASBHAC8(X23)
	h. Read aloud signs and labels	
<b>PATR</b> Parent's Attitudes Toward Reading	HQ11. Please indicate how much you agree with the following statements about reading	ASBHSTM1(X24)
	a. I read only if I have to	ASBHSTM2(X25)
	b. I like talking about books with other people	ASBHSTM3(X26)
	c. I like to spend my spare time reading	ASBHSTM4(X27)
	d. I read only if I need information	ASBHSTM5(X28)
	e. Reading is an important activity in my home	

Note: StQ = Student Questionnaire; HQ = Home Questionnaire.

$\Lambda$ :  $p \times q$  matrix of scalars called factor loadings. This matrix describes how the observed variables  $x$  are related to the latent factors  $f$ .

$\mu$ :  $p \times 1$  vector of scalars. Most software assumes by default that  $\mu = 0$  and analyze centered  $x$  variables, i.e.,  $x - \bar{x}$ .

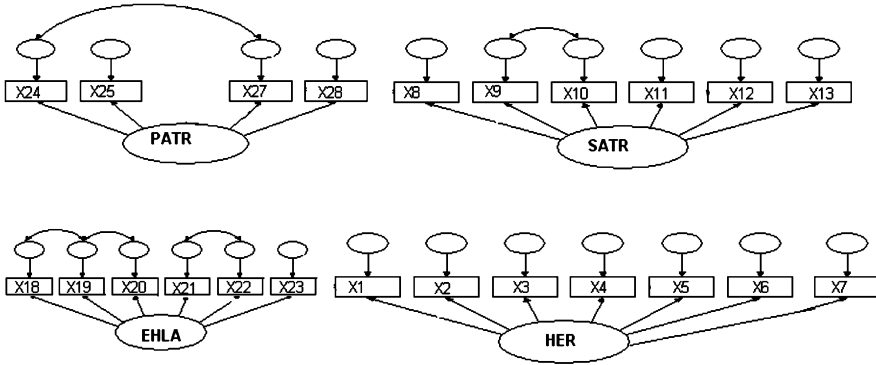


Fig. 1. Measurement models.

For example, we can write the following model for 'parents attitudes toward reading' latent variable PATR and indicators  $X_{24}, X_{25}, X_{27}, X_{28}$ :

$$\begin{pmatrix} X_{24} \\ X_{25} \\ X_{27} \\ X_{28} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} + \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \\ \lambda_{41} \end{pmatrix} PATR + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}$$

and the variance structure of  $PATR$  and  $\varepsilon$

$$var(PATR) = \phi_{11}, \quad Var(\varepsilon) = \begin{pmatrix} \psi_{11} & 0 & \psi_{13} & 0 \\ 0 & \psi_{22} & 0 & 0 \\ \psi_{31} & 0 & \psi_{33} & 0 \\ 0 & 0 & 0 & \psi_{44} \end{pmatrix}.$$

Other models can be expressed in a similar way.

Fig. 1 represents measurement models for home educational resources (HER), early home literacy activities (EHLA), students' attitudes towards reading (SATR), and parents' attitudes towards reading (PATR).

It should be noted that models slightly differ from proposed in an article by Hansen *et al.* [1]. In some cases, we introduced covariances among other pairs of residuals. For Lithuanian data, there was no need to introduce second residual factor to HER model. For the PATR, we saw that the indicator X26 ("I like to spend my spare time reading") caused high multicollinearity in the model, and we removed the index from it.

As we noted earlier there are statistically significant differences between the achievements of students from urban and rural schools in Lithuania. Therefore, we tested models for all data and two groups separately. In Table 2 results from the tests of models fit are reported. We should mention that group sizes mirror the proportions in the population from which the samples were drawn (1863 and 432 cases accordingly).

Table 2. Fit indices for measurements models

SAMPLE	Measurement model	Chi-square	df	RMSEA
TOTAL	HER	141.97	14	.06
	SATR	107.09	8	.069
	FHLA	24.513	6	.035
	PATR	.648	1	.000
URBAN	HER	76.836	14	.049
	SATR	74.638	8	.067
	FHLA	16.115	6	.030
	PATR	.340	1	.000
RURAL	HER	48.858	14	.076
	SATR	36.591	18	.091
	FHLA	10.291	6	.041
	PATR	.198	1	.000

RMSEA<sup>1</sup> is used as measure of fit. By convention, there is a good model fit if RMSEA  $\leq$  .05 (some researchers suggest RMSEA  $\leq$  .06 as the cutoff for a good model fit). There is adequate fit if RMSEA  $\leq$  .08. So we see, that all one-factor measurement models fitted the data quite well (except SATR model for rural data).

Analyzing factor loadings ( $\lambda$ 's) for measurement models, we obtained easily interpretable results. For example, in case of Early Home Literacy Activities (EHLA), larger loadings are for reading signs (.62) and books (.54), and the smallest, for singing songs (.22). In case of Home Educational Resources (HER), the largest loadings are for number of children and all books at home, as well as level of parents education (.74, .65, and .58, respectively), whereas possession of study table or newspaper at home had lowest factor loadings (.17 and .11, respectively). In case of Parents' and Students' Attitudes Towards Reading (PATR and SATR), the largest factor loadings were seen for the enjoyment of talking about books with others (.69) and enjoyment of reading (.85), respectively.

Finally, measurement models for each construct were gathered in one full Structural model (see Fig. 2). Correlations were estimated between latent variables and reading achievement. The results are presented in Table 3.

We see that the highest correlation with the Reading Achievement both in Sweden and Lithuania was in case of Home Educational Resources (HER). Comparing correlations for Swedish and Lithuanian data, we see that both HER and Parents Attitude Towards Reading (PATR) have similar correlations in the two countries, whereas Early Home Literacy Activities (EHLA) and especially Students' Attitude Towards Reading (SATR) correlate stronger with the Reading Achievements of Swedish students than Lithuanian ones.

It was of interest to test if the structural model would be the same for different school locality. Therefore, we tested the invariance of the model for urban and rural

<sup>1</sup>Root Mean Square Error of Approximation

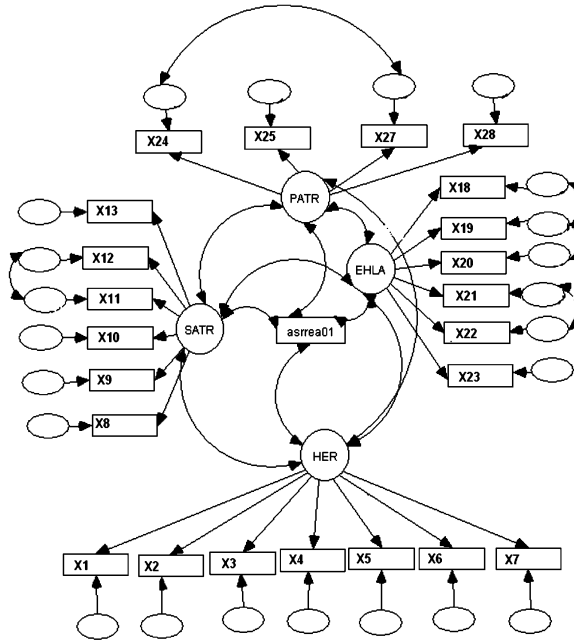


Fig. 2. Structural model for reading achievement.

Table 3. Correlation with reading achievements

Construct	SWE	LTU		
	Total	Total	Rural	Urban
HER	.46	.489	.461	.447
FHLA	.31	.248	.183	.279
PATR	.25	.350	.249	.228
SATR	.36	.198	.278	.201

schools' groups. Hypothesis about the invariance of the models is rejected at significance level .05 ( $\chi^2 = 34837, df = 239$ ).

Therefore, it is also of interest to compare correlations between the latent variables and Reading Achievement for the urban and rural schools in Lithuania. We see that urban and rural schools show similar correlations in case of HER and PATR, whereas there are differences in case of EHLA and SATR.

Urban group shows higher correlation between EHLA and Reading Achievement, whereas rural group shows higher correlation between SATR and Reading Achievements.

## Conclusions

The constructed models explain well reading achievement for Lithuania. The correlations of latent variables with the reading achievement were different in case of Sweden and Lithuania. Moreover, there were also differences in correlations of latent variables with the reading achievement for urban and rural schools within Lithuania. The highest correlation in Sweden and both urban and rural schools in Lithuania was seen between the Home Educational Resources and Reading Achievement.

## References

1. K.Y. Hansen *et al.*, Measures of self-reported reading resources, attitudes and activities based on latent variable modelling, in: *Proceedings of the IRC-2004*, vol. 3 (2004), pp. 65–79.
2. G.D. Israel *et al.*, The influence of family and community social capital on educational achievement, *Rural Sociology*, **66**(1), 43–68 (2001).
3. B. Jensen, A. Seltzer, Neighborhood and family effects in educational progress, *The Australian Economic Review*, **33**(1), 17–31 (2000).
4. M.O. Thirunarayanan, The “Significantly Worse” phenomenon: a study of student achievement in different content areas by school location, *Education and Urban Society*, **36**(4), 467–481 (2004).
5. J.R. Campbell *et al.*, *Framework and Specifications for PIRLS Assessment 2001*, International Study Center, Boston College (2001).
6. A. Eljio, Impact of home factors on students’ reading achievements, *Liet. matem. rink.*, **44** (spec. nr.) (2004).
7. *Database of PIRLS 2001*  
[www.pirls.com](http://www.pirls.com).

## REZIUMĖ

### **A. Eljio, G. Murauskas. Lietuvos mokinių skaitymo gebėjimų struktūrinė analizė**

Įvairių faktorių įtaka mokinių edukaciniams pasiekimams jau seniai nagrinėjama mokslininkų. Šiame straipsnyje, naudojant tarptautinio skaitymo gebėjimų tyrimo PIRLS 2001 duomenis, analizuojamas mokinių skaitymo gebėjimų struktūrinis modelis. Hansen *et al.* [1] pasiūlyti matavimo modeliai, susiejantys gebėjimų latentinius faktorius ir jų indikatorius, buvo pritaikyti lietuviškiems duomenims. Sudarytas pilnas struktūrinis modelis, aprašantis skaitymo gebėjimų ryšį su latentiniais faktoriais. Struktūrinė analizė atlikta AMOS 5.0 programa. Parodyta, kad skaitymo gebėjimų ir latentinių faktorių koreliacijos skiriasi tiek tarp švedų ir lietuvių duomenų, tiek tarp Lietuvos kaimo ir miesto duomenų.