Modelling TIMSS Data in a European Comparative Perspective: Exploring Influencing Factors on Achievement in Mathematics in Grade 8 *

Klaas Bos and Wilmad Kuiper
University of Twente, Enschede, the Netherlands

ABSTRACT

In international comparative studies like TIMSS data analysis is aimed at differences and similarities among education systems (countries). In this article the outcomes are presented of explorative path analysis on data collected with grade 8 students and classrooms in eight Western and two Central European education systems. For the 10 education systems the resulting general path model explains 19% or less of the variance in achievement in mathematics. In many systems home educational background and students’ attitude towards mathematics have a positive relation with achievement in mathematics, out-of-school activities a negative. Due to the psychometric quality of scales and non-availability of measures of important factors at classroom level (e.g., time on task and teacher’s expectation), no significant results were found of factors that can be manipulated by policy makers.

INTRODUCTION

In many countries education has high priority for policy makers. Children have to learn basic knowledge and skills they need in future life and in their professional careers. Reading, mathematics and science are major

*We are very grateful to our TIMSS colleagues and their staff members of the nine countries for allowing us to use their data and for reading and commenting on an early draft of this paper: Christiane Brusselmans-Dehairs (Belgium-Flemish), Cristian Monseur (Belgium-Flemish), Wendy Keys (England), Svein Lie (Norway), Ingemar Wedman (Sweden), Peter Weng (Denmark), Algirdas Zabulionis (Lithuania), Jana Strakova (Czech Republic), and Jürgen Baumert (Germany).

Address correspondence to: Klaas Bos, University of Twente, Faculty of Educational Science and Technology, Department of Curriculum, P.O. Box 217, 7500 AE Enschede, the Netherlands. Tel.: +31 53 4893955. Fax: +31 53 4893759. E-mail: boskt@edte.utwente.nl

Manuscript submitted: March 30, 1998
Accepted for publication: January 19, 1999
subjects. In every country questions can be raised about the quality and effectiveness of their education in terms of achievement results of students. These questions can be addressed in international comparative studies like TIMSS. The results of the initial data analyses carried out by the National Center of all participating countries and the International Study Center of TIMSS in Boston consists of descriptive data (e.g., Beaton et al., 1996). In order to find more comprehensive results, relational analyses on data of two or three different levels (i.e., student, teacher/school and country) are necessary. In this article the results of secondary data analyses on the mathematics data in grade 8 (population 2) from 10 European education systems are presented. The 10 education systems are: Belgium-Flemish, Belgium-French, Czech Republic, Denmark, England, Germany, Lithuania, Norway, Sweden, and the Netherlands. Participants are referred to as ‘education systems’ instead of ‘countries’. The reason for using this term is that it allows for the distinction between Belgium-Flemish and Belgium-French. By means of explorative data analyses a general path model was developed.

**RESEARCH QUESTION**

The major goal of this secondary analysis study on TIMSS data carried out by the Dutch National TIMSS Center is to find relations between achievement in mathematics and constructs (factors) at student and teacher levels. The main research question to be answered is: *To what extent can variances in the overall mathematics score for grade 8 students in 10 European education systems be explained by variances in the scores on constructs at student and teacher levels and to what extent are these outcomes generalisable across education systems?*

This question can also be formulated as: What can be learned about mathematics achievement of grade 8 students, and the factors at student and classroom levels that may be associated with that achievement across 10 education systems?

In answering this question the developmental process of the TIMSS instruments, in particular the questionnaires, must be taken into account. The developmental process of the TIMSS questionnaires has been documented by Schmidt and Cogan (1996, pp. 5-1 – 5-13). Many variables (being individual items or subsets of items) were included in the draft versions of the TIMSS questionnaires after discussions between participating countries. The bottom line for these discussions was the IEA research model (Robitaille & Garden, 1996) consisting of three curriculum
levels (intended, implemented, and attained), and three levels in a school (school, class, and student). This research model has been transformed into the TIMSS conceptual framework called ‘The Educational Experience Opportunity’ (Schmidt & Cogan, 1996, pp. 5–8, Fig. 5–4). Below, this conceptual framework for TIMSS is described. Nevertheless, considering the TIMSS questionnaires it is not very clear which important factors have been operationalised. They do not contain well-tested scales necessary to operationalise all important constructs. This is the reason why the data analysis that was carried out is ‘secondary’ and also explorative in nature. At this place a major implication of this fact must be stressed. Secondary explorative data analysis, including path and scale analysis (principal component analysis and reliability analysis), can result in the conclusion that some predictor variables appearing to be important in relation to mathematics achievement from the literature could only be partially covered or not covered at all by the TIMSS questionnaires. Such limitations of the available data with respect to the research question must be taken for granted. Considering the design of the study in all countries another important limitation of the available data is that (due to a restricted financial budget and to avoid overburning of schools) only one class per grade per school was investigated. Thus, school effects cannot be distinguished from class effects. The first data explorations will be carried out at one level. Teacher data will be disaggregated to student level. Below, the latter will be explained further.

CONCEPTUAL FRAMEWORK DATA EXPLORATION

Conceptual Framework for TIMSS
The conceptual framework for TIMSS was derived from earlier IEA studies and from the literature on educational indicators. It was based, in a fundamental way, on the conceptual framework for the Second International Mathematics Study (SIMS) (Robitaille & Maxwell, 1996). In this framework a distinction was made between mathematics curricula at three levels (cf. the editorial of this special issue): the intended curriculum as transmitted by national or system level authorities (that which a society would like to have taught), the implemented curriculum as interpreted and translated by teachers according to their experience and beliefs for particular classes (that which is actually taught) and the attained curriculum as that part of the intended curriculum learned by students which is manifested in their achievement and attitudes (that which students actually learn) (Travers, Garden, & Rosier, 1989, pp. 4–5). The concepts of the intended,
the implemented and the attained curriculum were adopted by TIMSS (see Figure 1). The variables studied in this article can be located at the implemented (teacher variables measured by means of the teacher questionnaire) and the attained curriculum level (student background variables and student achievement in mathematics).

As stated, studying the TIMSS questionnaires, for many items and sets of items it is not very clear of which constructs (factors) they are operationalisations. The curriculum based conceptual framework for TIMSS (the three curriculum model) can be filled out by constructs at teacher (or implemented curriculum) level and student (or attained curriculum) level that potentially influence student achievement. In order to select and to explore these factors a basic conceptual model has been selected.

**Basic Conceptual Model for TIMSS Data Explorations**

The three curriculum model of TIMSS can be partly transformed into a model of educational effectiveness. Looking for factors that contribute to education in mathematics and that covary with achievement in mathematics, conceptual models of educational effectiveness have a lot to offer. Creemers (1994) developed a conceptual framework for education at classroom level: model of educational effectiveness (see Fig. 2). This framework was based on Carroll’s model of school learning (Carroll, 1963, 1989). Creemers’ work can be regarded as a review and summary of the empirical research on effective instruction. Together with the review of
Scheerens (1992) resulting in his model of school effectiveness, Creemers’ model can be seen as an extension and refinement of Edmonds’ five-factor model of (1) strong educational leadership, (2) emphasis on basic skills achievement, (3) safe and orderly climate, (4) high expectations of students’ achievement, and (5) frequent evaluation of students’ progress (Edmonds, 1979). In the models of Scheerens and Creemers four levels are distinguished: the country, the school, the classroom, and the student levels.

In relation to the main research question it is necessary to detect influencing factors on achievement in mathematics, particularly the factors at teacher level that can be manipulated by policy makers. Therefore, Creemers’ conceptual framework of educational effectiveness, including the lists of potentially effective factors at each of the four levels, is regarded as the basic model. Here, the basic model serves primarily as a classification model of potentially effective educational factors and not as a model that
will be tested. The possible links between the different blocks of variables (see the arrows in Figure 2) are premature. Not all of these links will be taken into account in the explorative path analyses applied on the data (see below). In this model, instruction, teacher and student characteristics are taken into account in particular. At the same time instruction, teacher and student characteristics are the main components of the major data explorations presented here.

In Figure 2 it can be seen that student achievement is influenced by effective learning time and by the opportunity to learn (OTL). Effective learning time consists of the time students are willing to spend on school learning and on educational tasks as well as of the time offered in the instructional process (Creemers, 1994). Several student background factors influence achievement in addition. Important student background factors are motivation, attitude towards school (in this case mathematics in particular) and socio-economic status. At teacher or classroom level Creemers distinguished three main components of quality of instruction: curriculum, grouping procedures and teacher behaviour. On the basis of former research he elaborated these components. Potentially effective features of the curricular materials are:

- explicitness and ordering of goals and content;
- structure and clarity of the content;
- advance organisers;
- material for evaluation of student outcomes, feedback and corrective instruction.

Creemers describes three effective characteristics of grouping procedures:

- mastery learning;
- ability grouping;
- co-operative learning.

Potentially effective characteristics of teacher behaviour on the basis of Creemers' study are:

- management and orderly and quiet atmosphere;
- homework;
- high expectations of student progress and outcomes;
- clear goal setting;
- structuring the content;
- clarity of presentation;
- questioning;
- immediate exercise after presentation of new content;
- evaluation, feedback, and corrective instruction.
The two ‘boxes’ at the top of Figure 2 refer to factors at country and school levels. Scheerens and Bosker (1997) reviewed school effectiveness studies. They assume an indirect influence of school level characteristics via class teaching techniques on student achievement, but also a direct influence from school characteristics is recognised. Examples of school variables are educational leadership, orderly and secure environment and high expectations of student progress. Some of the factors that “are considered to work in education” (Scheerens & Bosker, 1997) are the same as distinguished by Creemers (1994) at classroom level. Examples of these factors are effective learning time, classroom climate and evaluating student achievement. Creemers says about this list of effective school factors that “most of the factors are in fact reflections of the indicators of quality of instruction, time and opportunity to learn at classroom level” (Creemers, 1994, p.120). Creemers defined all school level factors in his model as conditions for classroom level factors. Thus, in his model only those school level factors were selected that are conditional for and directly related to quality of instruction, time or opportunity to learn. In the review of Scheerens and Bosker (1997), a list of school level factors is presented with a direct influence on student achievement (e.g., school climate).

As far as the factors belong uniquely to the ‘country level box’ or the ‘school level box’ they will not be taken into account in the data explorations presented in this article.

Selection of Factors Operationalised in TIMSS

The TIMSS data have been explored on the basis of the list of potentially effective educational factors at student and teacher levels. As stated, not all of these factors can be traced in the TIMSS student and teacher questionnaires. For example, many characteristics of curricular materials have not been measured via the teacher questionnaire. The factors for which indicators are available in the TIMSS instruments have been selected for further explorative analyses. In Table 1 the list of potentially effective factors from the basic conceptual framework (model of educational effectiveness, see Figure 2) and their indicators available in TIMSS are presented.

The indicator in TIMSS of ‘social background’ of the student is ‘educational level of mother and father’. In all countries the percentage of missing values was too high (more than 20%) to allow some kind of imputation to replace the missing values. Therefore, the factor ‘social background’ is indicated in TIMSS by a proxy variable: ‘number of books in the home of the student’. The bivariate Pearson product-moment correlation between this variable and ‘mathematics score’ varies from .14 to .34 in the 10 countries. The ‘number of books’ is called a proxy variable because the
TIMSS questionnaires do not contain a satisfying alternative for educational levels of mother and father as an indicator for social background. Considering Table 1 one can conclude that not all potentially effective educational factors distinguished by Creemers (1994) and by Scheerens and Bosker (1997) at student and classroom levels are available in the TIMSS instruments. The indicators available in the TIMSS questionnaires will be described in more detail in the section “Results”.

The path analysis was carried out only at student level. Therefore, some of the teacher or classroom variables have been disaggregated to student level. Those variables have been marked by an * in Table 1. The dependent variable to be explained is ‘achievement in mathematics in grade 8’. It is measured by means of the international mathematics TIMSS test. Each grade 8 student tested in each country got an international mathematics score based on the TIMSS test. The scores were standardised with a mean of 500 and a standard deviation of 100. In addition to the potentially effective factors from the basic model of educational effectiveness a few

Table 1. Available Indicators in TIMSS Questionnaires of Potentially Effective Educational Factors with regard to Achievement in Mathematics.

<table>
<thead>
<tr>
<th>Potentially effective educational factor</th>
<th>Indicator in TIMSS questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td></td>
</tr>
<tr>
<td>social background</td>
<td>educational level mother and father (proxy: number of books at home)</td>
</tr>
<tr>
<td>motivation</td>
<td>attitude towards mathematics</td>
</tr>
<tr>
<td></td>
<td>success attribution</td>
</tr>
<tr>
<td></td>
<td>maternal expectation; friends’ expectation</td>
</tr>
<tr>
<td>time on task (effective learning time)</td>
<td>number of minutes mathematics per week*</td>
</tr>
<tr>
<td>Classroom</td>
<td></td>
</tr>
<tr>
<td>Classroom management</td>
<td>teaching style: student oriented or teacher centered (student’s perception)</td>
</tr>
<tr>
<td>orderly and quiet atmosphere</td>
<td>class climate (student’s perception)</td>
</tr>
<tr>
<td>homework</td>
<td>kind of homework*</td>
</tr>
<tr>
<td>evaluation, feedback and corrective instruction</td>
<td>kind of decisions based on assessment outcomes*</td>
</tr>
<tr>
<td>grouping procedures</td>
<td>instructional formats: co-operative learning*</td>
</tr>
<tr>
<td>School</td>
<td></td>
</tr>
<tr>
<td>school climate</td>
<td>safety at school (student’s perception)</td>
</tr>
</tbody>
</table>

Note. The indicators marked by an * have been disaggregated from teacher to student level.

other variables at student and teacher/classroom levels were included in the data explorations: student’s gender, student’s out-of-school activities and class size.

DATA ANALYSIS

The data explorations aiming at factors that influence achievement in mathematics in the 10 education systems consist of scale and path analysis. Principal components analysis and reliability analysis have been carried out on sets of items referring to one factor. Sets of items with a reliability coefficient Cronbach \( \alpha \) of at least .50 in 4 or more of the 19 countries involved in the analyses have been selected as a composite variable indicating one potentially effective educational factor from Creemers’ model. The threshold of .50 is rather low. The reason for not setting the minimum on .60 or .70 is that in that case only a few scales (sets of items) will be left. The results of all scale analyses are described in Bos (in preparation). The selected factors and their indicators of the path model to be explored are shown in Figure 3 and described below. The multiple indicators (scales) of these factors meet the modest reliability requirement.

The basic conceptual framework for this study has been directed to the research question. Looking at both the framework and the question, detection of influencing factors on achievement in mathematics can be done by means of statistical path analyses. Each influencing factor selected can show a single correlation with achievement in mathematics. To be able to find the significant factors it is necessary to calculate the intercorrelations of each important factor. The relative contributions of the different factors at student and teacher/classroom levels to educational outcomes and their interrelationships must be assessed accurately. These different variables cannot only have a direct influence on the dependent variable ‘international mathematics score’ but also an indirect influence via one of the other independent variables.

In this secondary analysis on TIMSS data the PLSpath approach has been applied. ‘PLSpath’ stands for ‘Partial Least Squares path analysis’ technique (Sellin, 1990, 1992). Since the path model has been developed post hoc (decisions concerning instruments and associated variables were made before the model was developed), the nature of the analysis is seen as more exploratory than confirmatory. The PLS technique has been developed especially for research situations that require a great deal of exploratory analyses. Other approaches like LISREL and AMOS, on the other hand, were designed primarily for situations that require confirmato-
Applying PLS consists of two main steps:

1. Estimation of latent variables (constructs or factors) at student/classroom level as linear composites of their associated manifest variables (indicators, items from the TIMSS questionnaires) by means of either principal component analysis or by means of regression analysis. This is called the outer PLS model.

2. Estimation of the direction and strength (path coefficients) of links between latent variables. This estimation is done by means of ordinary least squares regression applied to each equation (endogenous latent variables predicted by two or more other latent variables) separately and results in the estimated recursive inner PLS model.

Because PLSpath does not provide a goodness-of-fit measure, there is no criterion or cut-off point available for making a distinction between adequate and non-adequate fit of the path model. Instead, to consider the value of the estimations of the links between latent variables a jackknife procedure can be applied. For more details of the PLSpath technique, publications of Janssen Reinen (1996), Sellin and Keeves (1994), Sellin (1990, 1992), Anderson, Ryan, and Shapiro (1989), and Wold (1982) are recommended.

RESULTS EXPLORATIVE PATH ANALYSIS

The most important question in this study is whether variances in mathematics achievement of grade 8 students within eight Western European countries and two Central European systems can be explained (or at least described) by means of different path coefficients existing in a general path model. ‘General’ means that the estimated relations between manifest and latent variables (the outer model) and between latent variables (the inner model) are explored for more than one education system. It would also have been possible to develop a different model for each system such that each model would explain more of the variance in the criterion mathematics achievement score than would one general model applied to each system in turn. The advantage of having a general model is the facilitation...
of the identification of the different influences in each of the systems. It is, for example, of interest to know that factor X has a strong effect (meaning a high path coefficient) on mathematics achievement in some systems but not in the other systems (cf. Zabulionis, 1997). Before describing the results of the PLS analyses the research group for each education system will be described.

**Education Systems and Achievement in Mathematics**

The definition of population 2 in TIMSS refers to the two adjacent grades containing the majority of the 13-year-old students. For 5 of the 10 countries this means grades 7 and 8 with respectively 7 and 8 years of schooling. In the analyses presented here, we focus on grade 8 only. The number of years of formal schooling for the three Scandinavian countries is not 8 but 7 years. In Denmark, Norway and Sweden students enter school 1 year later than in the other countries. Consequently, the majority of the 13-year-old students have 1 year of schooling less than the students of the same age in the other countries. In England the number of years of formal

### Table 2. Mean Score Mathematics TIMSS Test Grade 8 of Eight Western European Education Systems and Two Central European Systems.

<table>
<thead>
<tr>
<th>Education system</th>
<th>Number of Students</th>
<th>Number of Teachers (= classrooms)</th>
<th>Mathematics achievement TIMSS test (spring 1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean (sd)</td>
</tr>
<tr>
<td>BFL Belgium-Flemish</td>
<td>2722</td>
<td>144</td>
<td>565 (92)</td>
</tr>
<tr>
<td>NLD Netherlands</td>
<td>1814</td>
<td>88</td>
<td>541 (89)</td>
</tr>
<tr>
<td>BFR Belgium-French</td>
<td>1801</td>
<td>91</td>
<td>526 (86)</td>
</tr>
<tr>
<td>SWE Sweden</td>
<td>3296</td>
<td>193</td>
<td>519 (85)</td>
</tr>
<tr>
<td>GER Germany</td>
<td>1964</td>
<td>99</td>
<td>509 (90)</td>
</tr>
<tr>
<td>ENG England</td>
<td>812</td>
<td>109</td>
<td>506 (93)</td>
</tr>
<tr>
<td>NOR Norway</td>
<td>2411</td>
<td>106</td>
<td>503 (84)</td>
</tr>
<tr>
<td>DNK Denmark</td>
<td>1859</td>
<td>118</td>
<td>502 (84)</td>
</tr>
<tr>
<td>CSK Czech Republic</td>
<td>2876</td>
<td>130</td>
<td>564 (94)</td>
</tr>
<tr>
<td>LTU Lithuania</td>
<td>2080</td>
<td>123</td>
<td>477 (80)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>21635</strong></td>
<td><strong>1201</strong></td>
<td><strong>521 (87)</strong></td>
</tr>
</tbody>
</table>

*Note.* Mean (sd) = mean (standard deviation); SE = standard error (Source: Beaton et al., 1996; p.22, Table 1.1); the column ‘students’ contains the number of students that could be linked to the teachers as presented in the next column. In England, students were selected randomly from all grade 8 students in each school. The selected students in each school were, therefore, drawn from a number of mathematics classes. For this reason, the mean number of students per teacher was much lower in England than in the other countries (where intact classes were randomly selected).
K. BOS AND W. KUIPER

schooling is 9. The differences in ‘years of formal schooling’ can be taken into account when describing the PLS results.

In Table 2 for each education system the number of students and teachers included in the analyses are presented. In Beaton, Mullis, et al. (1996; Table 1.1, p.22) it can be read that the education systems under investigation differ in the mean achievement on 150 mathematics items of the international mathematics test. In the final columns of Table 2, the figures of the 10 education systems are presented.

The number of students involved in the analyses consists of the students that could be linked to a teacher (see the 2nd and 3rd column of Table 2). The abbreviations of the education systems used in the first column of Table 2 will be used continuously.

**A General Student/Classroom Path Model**

In Figure 2, the basic conceptual framework of educational effectiveness has been presented. The selected factors that potentially influence achievement in mathematics with multiple indicators that meet the reliability criterion (Cronbach $\alpha > .50$; see Table 1) were included in a general student/classroom path model as presented in Figure 3.

---

Fig. 3. Recursive student/classroom path model.
This model forms the basis for the first exploration of a path model by means of the two steps of PLS. The potentially influencing factors (latent variables) can be explored in step 1. In Figure 3 the possible links between the different variables at classroom level (disaggregated to student level) and student level are presented (variables at society/country- and school level from Figure 2 will not be taken into account in this exploration of a path model). The possible links can be introduced to the model once the set of estimated latent variables at student/classroom level has been determined. Next, these links can be explored in step 2 of the PLS analyses. The measurement model is a general one and is invariant across countries. The general model will be explored separately for each country in step 2 of the PLS analyses.

Step 1: Estimation of Latent Variables
Clustering of manifest variables should result in latent variables. Like Postlethwaite and Wiley (1992, p.125) write, it must be stressed that clustering of items should reflect meaningful homogeneity within clusters both conceptually and empirically. Conceptually means the latent variable must make sense, that is has a meaning in literature; empirically means the latent variable must have meaningful loadings on one factor in a principal component analysis and a correlation higher than 0.10 (absolute value) with the criterion variable. For example clustering items with respect to class climate (orderly and quiet atmosphere) can form a latent variable (construct or factor) if the meaning of each of these items can be linked to class climate (conceptual homogeneity) and the loadings of each of these items are high enough plus the correlation of the sumscore of the individual item scores with ‘mathematics achievement’ is higher than 0.10 (empirical homogeneity).

Because the general path model must apply not only to one sole education system but to all systems involved, the criteria for keeping a manifest variable have been set in advance. A manifest variable is an item of a TIMSS questionnaire or a composite variable (scale score of more than two items) which has been selected after scale analysis (see above). Meaningful loadings are loadings of at least .40. A manifest variable was kept in the set associated to 1 latent variable if in at least 4 out of the 10 systems its loading is .40 or higher (Campbell, 1996). The loadings of all manifest variables associated with a latent variable in Figure 3 conform to this requirement.

The bivariate pearson pm correlation between a manifest variable and the criterion variable ‘mathematics score TIMSS test’ is lower than .10 in more than four countries for maternal expectation, friends’ expectation,
class climate, teaching style (respectively student oriented and teacher centered), co-operative learning, homework and assessment. Following the ‘empirical’ rule stated above, these latent variables should be removed from the path model presented in Figure 3. Removal will not be done, because it concerns a great number of latent variables. The consequences of this ‘offence’ will become clear in step 2 of the PLS analysis.

The latent variables included in the path model of Figure 3, including those that are associated with only one manifest variable, can be described as follows on the basis of the final PLS outer model (see Figure 3, column-wise from left to right). The descriptions of some variables correspond to the descriptions given in another study on influencing factors on achievement in mathematics (for example ‘teaching style’). In this study, student data from TIMSS of nine Central and Eastern European countries were involved (Zabulionis, 1997).

**Homework**

‘Homework’ can be described as the kind of homework assigned by the teacher. Two composite variables serve as manifest variables for ‘homework’: the frequency of assigning homework primarily related to the textbook (four items) and the frequency of assigning homework that can be characterized as ‘applied schoolwork’ (six items, e.g., small investigations and working individually on long term projects). A high score means the students have been assigned often both homework related to the textbook and ‘applied’ homework. ‘Homework’ is a disaggregated teacher factor.

**Teaching style**

This construct can be defined by using several questions of the student questionnaire. Eight of these variables reflect a more student oriented teaching style and four other variables reflect a more teacher centered teaching style. Variables reflecting ‘Student oriented teaching style’ are: students work from worksheet on their own, students work on a mathematics project, students work in pairs or in small groups, students use daily problems when problem solving, check each other’s homework, discuss practical problems, students are asked what they know related to a new topic, solve an example related to a new topic. The four variables that reflect ‘Teacher centered teaching style’ are: teacher shows how to do mathematics problems, student copies notes from the board, teacher explains rules, and student follows textbook when teacher teaches.

All scores of these eight respectively four variables were inverted. The inverted scores mean: the higher the score the more it can be regarded as an indicator for student oriented respectively teacher centered teaching.
style. In order to get a meaningful score for ‘teaching style’, the mean score of the eight ‘student’ scores and the mean score of the four ‘teacher’ scores were calculated. The ‘student’ mean was subtracted from the ‘teacher’ mean. As a consequence, a high positive score on ‘teaching style’ means the class teaching was ‘more teacher centered’ and a negative score means the class teaching was ‘more student oriented’.

**School climate**
This construct consists of perceptions of the student about wrong behaviour of himself and other students (not only of their own classrooms but also of other classrooms) during the last month: something stolen and someone hurt. The raw scores of these variables (never, once or twice, three or four times and more than four times) were coded from high to low (never = 4, etc.). Hence, a high score means the school climate was safe and a low score means a poor school climate in the perception of the student.

**Student’s gender**
The values of boys and girls were respectively ‘2’ and ‘1’. Thus, a positive link between ‘gender’ and another variable means boys do ‘better’ or ‘more’ than girls.

**Maternal expectation**
This construct reflects the perception of the student of the extent to which his mother thought it is important for him to do well at school in mathematics, science and mother-tongue. A high score means the student perceives a great pressure from his mother to do well at school.

**Friends’ expectation**
‘Friends’ expectation’ has basically the same contents as ‘maternal expectation’, only it is the student’s perception of the expectation of friends.

**Success attribution mathematics**
Students were asked about four manifest variables to what extent they think it is needed to do well in mathematics: a lot of talent, to have good luck, to undertake lots of work, hard studying at home and to memorise the textbooks. The scores on the first and the third variable were inverted. Thus, a high score on the construct means the student thinks to do well in mathematics has mainly to do with talent and effort and not with good luck or memorising. A low score means the opposite, the student thinks to do well in mathematics is more a matter of good luck and memorising than a matter of effort.
Instructional formats
The extent to which students work in pairs or small groups according to the teacher is taken as the manifest variable reflecting co-operative learning. Co-operative learning is regarded as a potentially effective instructional format. This teacher variable has been disaggregated to student level.

Mathematics lesson climate
The construct ‘mathematics lesson climate’ is a perceptual measure. Students were asked for their perception of the climate during mathematics lessons. Three manifest variables (statements) reflect this construct: ‘students often neglect their schoolwork’ (scores were inverted to mean that ‘students did not neglect schoolwork but took it seriously’); ‘students are orderly and quiet’; ‘students do exactly as the teacher said’. A high score means the students perceived an orderly and quiet atmosphere during mathematics lessons.

Attitude towards mathematics
The attitude towards mathematics can be regarded as a predictor for achievement in mathematics but also as an independent variable. The TIMSS student questionnaire contains 10 questions (manifest variables) that potentially refer to attitude. All variables were recoded in such a way that a high score means a positive attitude towards mathematics and a low score means a negative attitude. Five manifest variables refer to liking mathematics, the other five refer to ‘the perceived importance of mathematics by the student’. A high score on attitude means the student likes mathematics and thinks mathematics is important for himself.

Home educational background
This construct reflects the educational level of the student’s home. It consists of one manifest variable: the number of books in the home. This is a proxy indicator for home educational background. As stated, in the TIMSS student questionnaire another indicator of ‘social background’ of the student was available: ‘educational level of mother and father’. However, in all countries the percentage of missing values was too high (more than 20%) to allow some kind of imputation to replace the missing values. Therefore, the factor ‘social background’ is indicated in TIMSS by a proxy variable: ‘number of books in the student’s home’.

Class size
Class size is the total number of students (girls + boys) in the classroom. The teacher has given these figures.
Effective learning time: Total number of minutes mathematics per week
This construct is called ‘mathematics instruction time’ and is equal to the total number of minutes the students get mathematics per week. Certainly, it is not the best indicator for effective learning time. A better indicator is not available in the TIMSS data set.

Assessment (evaluation, feedback, and corrective instruction)
This construct consists of one composite being the sumscore of three manifest variables dealing with the frequency of usage of assessment information for making decisions concerning a student’s school career:
• providing students’ grades or marks;
• report to parents;
• assign students to different programs or tracks.
A high score on the construct means the teacher uses assessment information often for making decisions with regard to a student’s school career. The scores of this teacher variable have been disaggregated to student level.

Out-of-school activities
Students can do a lot of out-of-school activities like having a paid job, watching tv or videos, and reading a book for enjoyment. Two composites refer to this construct:
• job-related: working for a paid job and doing jobs in the home;
• leisure time: being with friends, watching tv or video and playing computer games.
A high score on this construct means the student spends a lot of time on these activities.

Step 2: Exploration of the Links Between Latent Variables
In the second step of the PLS analysis the links between the latent variables (factors or constructs) can be drawn in many ways. The basic model of educational effectiveness (Fig. 2) served as the starting point for the exploration of possible links. The possible links have constraints in the sense that they are ‘recursive’. This means that when one latent variable is presumed to ‘influence’ another (i.e., if the second is an outcome of the first) then there can be no ‘influence’ of the second on the first. Besides, ‘circular’ relations are not allowed: if construct A affects construct B and construct B affects construct C, then construct C cannot affect construct A (cf. Zabulionis, 1997). Keeping this in mind, a set of direct and indirect relationships between latent variables was proposed (see Fig. 3). Achievement in mathematics is supposed to be influenced directly by out-of-school
activities, home educational background, student’s attitude towards mathematics, class climate during mathematics lessons, instructional formats (co-operative learning), effective learning time and usage of assessment results. Furthermore, from Figure 3 it can be seen that class climate and student’s attitude towards mathematics are supposed to be influenced by a number of latent variables. Homework, teaching style, school climate, friends’ expectations and attitudes are directly linked to class climate. Home educational background, student’s gender, maternal expectation, friends’ expectations and success attribution are directly linked to student’s attitude towards mathematics. Class size is supposed to influence effective learning time (together with class climate) and instructional formats.

The links between the latent variables in the general path model from Figure 3 have been estimated separately for each education system. The differences between path coefficients of the education systems can be described as follows. First of all, the percentage of variance in students’ mathematics scores explained by the latent variables of the path model is not higher than 19% (in England). In the other systems the $R^2$ varies from .07 (Denmark) to .18 (Netherlands). These results are partly the consequence of keeping some latent variables in the path model although they showed low correlations ($r < .10$) with achievement in mathematics. Generally, the path coefficients are not high. In most of the 10 systems 3 factors have significant influence on achievement (direct links): home educational background, out-of-school activities and attitude towards mathematics. Home educational background shows the highest (positive) path coefficients in most of the systems ($>.21$ in nine systems) together with out-of-school activities ($<-.10$ in seven systems). The path coefficient of out-of-school activities is negative, which means that the more time a student spends on jobs and watching television and playing games the less his or her achievement in mathematics is. Only in Germany, Norway, and Denmark this factor does not have a significant path to achievement. Attitude has a positive relation with achievement in 8 of the 10 systems (not in Germany and England). The path coefficient ‘attitude → achievement’ varies from .12 in Lithuania to .22 in Norway. The other four factors that were supposed to have a direct influence on achievement do not show a significant path coefficient in the majority of the education systems:

- class climate as perceived by the students (England is the exception with a significant path coefficient of .15);
- assessment usage;
- instructional formats (co-operative learning);
- effective learning time.
The latter three factors are disaggregated teacher variables. Disaggregating teacher data from teacher to student level might be one explanation for not finding significant path coefficients from these factors towards achievement in mathematics.

The direct links between attitude towards mathematics and five of the other latent variables do not differ very much with regard to significant path coefficients. The percentage of variance in attitude explained by the five latent variables is relatively high in England (25%) and in Norway (26%). In the other systems this percentage lies between 13% and 19%. In all 10 systems home educational background has no direct link to attitude. With a few exceptions the other four factors (student’s gender, maternal expectation, friends’ expectations, and success attribution) have a positive link to attitude. Again, the path coefficients are not very high (varying from .11 to .28). In some systems boys have a more positive attitude towards mathematics than girls. For Denmark, Czech Republic, and Lithuania the model shows no significant path coefficient between gender and attitude. The factor friends’ expectations of achievement influences attitude in nine systems (not in Denmark) as does maternal expectation in all systems.

The importance of hard working to do well in mathematics (success attribution) is positively related to attitude in all systems except in the Netherlands, Germany, and Lithuania. In seven systems the attitude of students who think hard working is important to do well in mathematics is more positive (they like mathematics more and they think mathematics is important for themselves) than the attitude of students who think hard working is not so important. This relation sounds theoretically logical. Therefore, the exception of the three systems mentioned is difficult to explain.

The percentage of variance in class climate (as perceived by the students) explained by five latent variables (homework, teaching style, school climate, friends’ expectations and student’s attitude) is 18% in Sweden. In the other systems this percentage is lower than 12%. Homework does not influence class climate in all systems and dominating teaching style does not influence class climate in the majority of the systems. The path coefficient from dominating teaching style to class climate is significant for four systems (Germany, Netherlands, Belgium-French, and Lithuania) and is not higher than .14. Student’s attitude towards mathematics is linked to class climate significantly in six systems with path coefficients varying from .09 to .19. The path coefficients for safety at school (nine systems, this indicator has not been measured in England) and friends’ expectations (nine systems) are of the same size.
With respect to two endogenous latent variables, effective learning time and instructional formats, the $R^2$ is very low for both variables ($<.07$). The accompanying path coefficients do not mean very much. The disaggregation of the teacher scores on effective learning time and instructional formats to student level decreases the possibility of finding meaningful links.

CONCLUSIONS AND DISCUSSION

The main research question was: What can be learned about mathematics achievement of grade 8 students and the factors at student and classroom levels that may be associated with it across countries? On the basis of the PLS outcomes, some preliminary conclusions can be drawn. A few factors, all at student level, seem to be important to explain variances in achievement in mathematics directly at least in six of the education systems: home educational background (positive relation), out-of-school activities (negative relation) and attitude towards mathematics (positive relation). With regard to attitude the factors student’s gender, maternal and friends’ expectations, and success attribution are important in most of the 10 systems. In general, the majority of the path coefficients do not exceed the value of .20 and the percentage of variance explained by latent variables of achievement in mathematics is lower than 20%. In the student path model explored by Zabulionis (1997), the $R^2$ in nine Central and Eastern European countries varied from .19 to .33. In the model he explored no teacher variables were involved, only variables measured at student level.

Considering the low $R^2$ of the three endogenous latent variables class climate, effective learning time, and instructional formats plus the low or non-significant path coefficients from these latent variables towards achievement in mathematics, one can conclude that these factors cannot be kept in the model unless better indicators (manifest variables) can be found in the TIMSS data sets. Besides, changing the links between the latent variables possibly can result in more meaningful paths. The latter is doubtful as long as the latent variables are reflected by the same manifest variables.

The factors indicated by disaggregated teacher variables (like homework, assessment, and co-operative learning) do not show convincing direct or indirect relations with achievement in mathematics. This result is generalisable across all education systems. At the same time, these factors are the factors that can be manipulated by policy makers with the knowledge in mind of the direct effects of factors at student level, like for
example student’s attitude, gender and success attribution. Knowledge of
the positive relationship between student’s attitude towards mathematics
and achievement in mathematics and between success attribution, gender
and attitude is useful in combination with knowledge about relationships
between these student and classroom factors like homework and assess-
ment (evaluation). The latter two can potentially be manipulated by policy
makers, subject departments or schoolboards. From the analyses shown
only covariates of these classroom factors can be found.

What can be done in future research to find more meaningful relations
between factors at classroom level and achievement in mathematics? First-
ly, comparative educational research needs more latent variables that can
be estimated by composites that are more reliable (Cronbach \( \alpha \geq .70 \)). The
data explorations presented in this article show for many composites in
many education systems a reliability coefficient not higher than .50. Nev-
evertheless, they were kept in the model. Above, it was already stated that
the instruments developed in TIMSS do not contain well-tested scales
necessary to operationalise all important constructs. Considering the TIMSS
questionnaires it was not very clear which important factors have been
operationalised and what their psychometric characteristics are. After the
data explorations some scales turned out to be reliable (like attitude and
maternal and friends’ expectations), yet the biggest problem with the
TIMSS data was the low reliability and the rather unknown construct
validity of scales that should be the operationalisation of potentially effec-
tive factors on achievement like effective learning time, homework, eval-
uation, and opportunity to learn. The latter is very important in many IEA
studies (with their three curriculum model, see Figure 1), but has not been
included into the path model presented in this article. The TIMSS database
does not contain convenient data for this variable. The only possibility is a
measure of ‘content coverage’. However, this is a rather crude measure for
opportunity to learn because it tells something about topics that have been
taught before the administration of the TIMSS test, but it tells nothing
about the coverage of the contents of the separate test items. Moreover, the
‘content coverage’ data is not very accessible to compute a composite with
a meaningful score.

Secondly, the disaggregation of scores on teacher variables should be
avoided. The design of international comparative studies like TIMSS should
be appropriate for multi-level data analysis (Bryk & Raudenbush, 1992).
In the data explorations presented here, the bivariate pearson correlations
between the output variable ‘mathematics score TIMSS test’ and the fac-
tors that were supposed to influence achievement directly (e.g., effective
learning time) were low (< .10) in many education systems. Disaggrega-
tion of teacher scores to student scores of some variables can lower their variance and therefore their correlations with achievement decreases. Factors at school and country levels which have not been used in the data explorations presented in this article, could be included in a multi-level model. Besides, student effects can be distinguished from classroom and country effects if a multi-level study design is applied. But again, a hierarchical linear model also needs well developed reliable and valid data to reach more useful results than could be reached by analysing a one level path model by means of PLSpath.

In future international comparative studies on student achievement and its influencing factors it is worthwhile to develop a conceptual model with factors that can both influence achievement and be manipulated by policy makers. Next, it is necessary to operationalise the selected factors in a reliable and valid way. Before carrying out the data collection agreement should be reached among participating countries about the conceptual model and, on the basis of pilot testing of questionnaire items, the operationalisation of its components. Many countries should be involved in this activity in order to get operationalisations that are meaningful, reliable and valid in many countries. The results of the data explorations presented in this article and more in-depth analyses on TIMSS data (to be done) can serve as the basis of such an activity. Above all, it is still worth trying to find out why some countries perform better than other countries on the same international achievement test and to learn from each other.

REFERENCES


Bos, K.Tj. (in preparation). Design and instruments of international comparative educational studies: The TIMSS case of Mathematics in Dutch lower secondary education in an European perspective. Enschede: University Twente, OCTO.


