The contribution of TIMSS to theories of educational effectiveness: a systematic review of the literature

M. Drent, University of Twente, marjolein.drent@utwente.nl
M.R.M. Meelissen, University of Twente, m.r.m.meelissen@utwente.nl
F.M. van der Kleij, University of Twente

Keywords: TIMSS, systematic review, educational effectiveness, school effectiveness.

1. Introduction

Since the first international comparative study on students' achievement in the beginning of 1960's, the interest of countries from all over the world to participate in studies initiated by the *International Association for the Evaluation of Educational Achievement* (IEA), has been growing rapidly. In the *First International Mathematics Study* (FIMS, 1964) just 12 countries participated. In TIMSS-2011 and PIRLS-2011 almost 70 countries are participating in one or both studies.

Participating in these studies is quite an enterprise in terms of costs, time and effort. Researchers from all participating countries work together in developing the achievement tests and context questionnaires (student, teacher school and curriculum questionnaires). The research instruments are translated and adapted for each country, and a lot of effort is put in to guarantee the international validity of these instruments. In each country, at least 4500 students from a representative sample of schools are tested, and these students, their teachers and schools principals are asked to fill in questionnaires about the context of teaching and learning. Each country has to make sure that the conditions for the assessment are the same. Students' answers on the open ended test items (about half of the test) are individually scored by trained scorers, and part of the items are doubled scored to ensure the reliability of the scoring. Finally, the data of all participating countries are collected, cleaned, weights are calculated and items response analyses are conducted for the achievement scores.

In return, TIMSS and PIRLS offer more than ranking countries in students' achievement in mathematics, science or reading. These studies provide an in-depth insight in the diversity in

education systems from all over the world. Furthermore, countries which participated in more than one of the TIMSS or PIRLS projects, are able to analyse trend data across assessments. Most importantly, these studies have the potential to contribute to effective education policy making (Mullis, Martin, & Foy, 2008; Postlethwaite, 1995; Postlethwaite & Ross, 1992). The studies are classroom based, which means that the context of teaching and learning can be related to students' achievement.

Increasing the knowledge about and the empirical support for what is effective in education, was one of the main reasons for the foundation of IEA in 1958. The founders viewed the world as "a natural educational laboratory, where different school systems experiment in different ways to obtain optimal results in the education of their youth" (http://www.iea.nl). By sharing the outcomes and characteristics of these 'experiments', international comparative studies offer empirical based insights in what matters in education within different contexts. For this, the data of these studies are well documented and made easy accessible for all who are interested.

Of all the different IEA studies conducted, TIMSS is the largest international comparative study in terms of longitude, participation and research populations. However, based on reviews of studies on TIMSS-1995 and -1999, Beaton and Robitaille (2002) concluded that the impact of TIMSS on educational effectiveness was still limited:

> A great deal of what have been published to date about TIMSS has focused on the achievement results, and on the ranking of the participating countries in "league tables" that have been published in wide variety of publications. Much less has been written and published so far about the more important–at least from an educational perspective–questions relating to ways and means of improving the teaching and learning experience in school mathematics and science for students in all countries .

It is not only the ambition of IEA and TIMSS that the project should contribute to the improvement of teaching and learning world wide. From the field of Educational Effectiveness Research (EER), there seem to be a growing interest to use international comparative studies to further develop their so called integrated multilevel educational effectiveness models (Kyriakides, 2006). The basic assumptions of these models are a) that influences on student achievement are multilevel, b) that the effect of manipulative factors

(manipulative by teachers, schools or policymakers) at the classroom and school level on students achievement, need to be controlled for student background characteristics, and c) that the higher levels provide the conditions for what happens in the classroom. According to Kyriakides (2006) there are two reasons why an international component should be added to EER. First, it is too simple to assume that what works in one education system will also be effective in another education system. Insights in the influence of contextual characteristics on effective education could prevent the simplistic transportation of knowledge from one education system to another. Secondly, international studies offer the possibility to study a full range of variation in school and classroom characteristics. This last argument is in line with IEA's laboratory view. The TIMSS-data offers the possibility to analyse the effects of manipulative school and classroom factors on students' achievement in mathematics and science, controlled for student characteristics within a wide range of school and country contexts.

2. Purpose and research questions

Recent contributions in academic journals (including special issues), reports and books on TIMSS as well as PhD studies on TIMSS and PIRLS data, indicate that researchers are becoming increasingly interested in analyzing the relation between teaching and learning characteristics and students' achievement in international contexts (Bos, 2002; Diepen, 2007; Howie, 2002; Kaya & Rice, 2009; Martin, Mullis, Gregory, Hoyle, & Shen, 2000; Meelissen & Luyten, 2008; Papanastasiou, 2008; Van den Broeck, Opdenakker, & Van Damme, 2005). With so many data collected since TIMSS-1999 (TIMSS-2003, TIMSS-2007, TIMSS-Advanced 2008), in so many countries, and with so many researchers and policy makers showing an interest in TIMSS, it seems likely that the contribution of TIMSS to theories on educational effectiveness has increased since the last review of TIMSS-studies by Beaton and Robitaille (2002). Therefore, the aim of this study is to give an overview of studies based on TIMSS data which contribute to theories of Educational Effectiveness Research. This is done by systematically analyzing these studies as they are represented in scientific journals since the release of the TIMSS-1995 data. Each article is carefully examined with the following research questions in mind:

1. *To what extent have secondary analyses on TIMSS data contributed to theories of educational effectiveness, in terms of the number and impact of these studies?*

2. *How can the TIMSS studies which contribute to theories of educational effectiveness be characterized and what is their quality?*
3. *To what extent have secondary analyses on TIMSS data contributed to theories of educational effectiveness in terms of the extent in which the findings of these studies support existing theories on educational effectiveness?*

The studies under review are publications on TIMSS data aimed at the relation between students' achievement and characteristics of teaching and learning on school, class and student level. In this study, educational effectiveness refers to manipulative (changeable) factors which are presumed to have a positive effect on students' achievement. For this, the study used the conceptual framework for school effectiveness by Scheerens (1997), which is developed within the tradition of EER. This framework has been used for several meta-analyses on effectiveness enhancing factors in education (Scheerens, Luyten, Steen, & Luyten-de Thouars, 2007). In the next section, a further explanation of this framework is given.

3. Conceptual Framework

The aim of school effectiveness research is to analyse the association of hypothetical effectiveness enhancing conditions of schooling and output measures such as student achievement. According to Scheerens et al. (2007) the idea of quality inherent in integrated school effectiveness models is that (pp. 55-56):

1) Outputs are the basic criteria to judge educational quality.
2) These output measures should be controlled for student characteristics in order to assess the added-value of schooling
3) Indicators for process characteristics should be selected based on correlations found with relatively high output. These measures should also be adjusted for the influence of student characteristics.
4) The model is multi-level, uniting effectiveness enhancing conditions at system, school, classroom, and individual student level (Scheerens & Creemers, 1989).

The basic model from system theory is often used as a conceptual framework to summarise the results of school effectiveness research (Scheerens, et al., 1997; Scheerens, et al., 2007; Scheerens, Seidel, Witziers, Hendriks, & Doornekamp, 2005). In this model (see Figure 1) the school as organizational unit is the starting point. A distinction is made between student (prior

achievement or SES), context (country or school characteristics such as size or student-body composition), input (teacher experience or parent support), process (educational leadership or opportunity to learn) and output characteristics (achievement). Only the characteristics on the process level (either classroom or school) can be directly influenced by the school and the teacher.

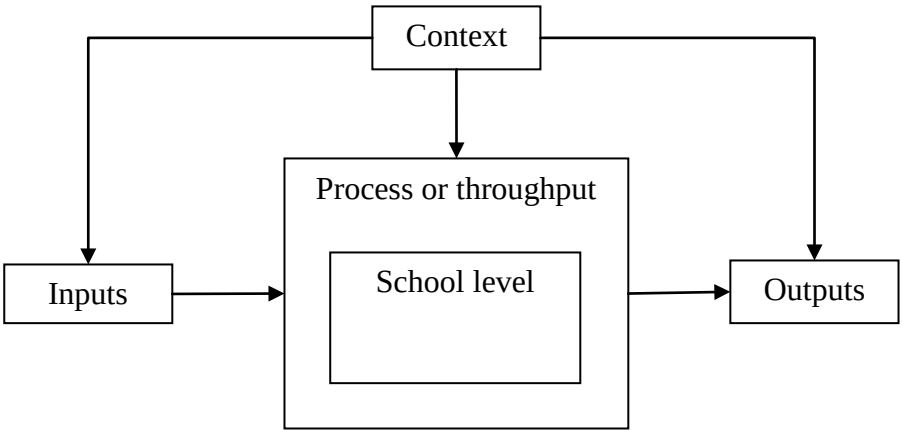

Figure 1. *A* b*asic systems model on the functioning of education (Scheerens, et al., 2007, p. 28)*

Based on systematic reviews and meta-analyses on studies conducted in developed countries, in which the output factor is achievement in core subjects, Scheerens et al. (2007) identified 12 school effectiveness factors and 5 instructional effectiveness dimensions (see Table 1 and Table 2). Table 1 also contains factors that could be defined at both school and classroom level. These factors are considered to be aggregated classroom level factors; for example, school policies stimulating a student oriented teaching approach.

Table 1 Effectiveness factors on school level (including aggregated classroom level factors), based on Scheerens et al. (2007, p. 62)

1. achievement orientation/high expectations/teacher expectations
2. educational leadership
3. consensus and cohesion among staff
4. curriculum quality/opportunity to learn
5. school climate
6. evaluative potential
7. parental involvement
8. classroom climate
9. effective learning time (classroom management)
10. structured instruction
11. differentiation, adaptive instruction
12. feedback and reinforcement

Table 2 Dimensions and effectiveness factors on classroom level, based on Scheerens et al. (2007, p. 194)

1. curricular dimension: opportunity to learn, strategies to learn about the deep structure of domain specific knowledge, and textbooks;
2. dimension of teacher orchestrated classroom management and climate creation: including time, achievement orientation, high expectations, disciplinary climate, activating measures, such as variation in representation formats, media, forms of practice, variation in applications (theoretical and authentic) grouping forms and differentiation/adaptive teaching;
3. teaching strategy dimension with two main sub-categories:
a) structured, direct teaching, mastery learning orientation, drill and practice;
b) constructivist oriented teaching strategy, teaching meta-cognitive strategies, cognitive activation, frequent open learning tasks, discovery learning, fading from more structured to more open assignments
4. climate dimension: support and positive interactions
5. evaluation and feedback dimension

Based on the school effectiveness framework and the four criteria of Scheerens et al. (2007) mentioned before, only studies which focus on manipulative process factors on school and classroom level (such as the factors mentioned above) related to mathematics or science achievement are included in this study. The choice to focus on process factors also means that manipulative factors on system level –for example the influence of central exit examinations (Jurges, Buchel, & Schneider, 2005)– are not part of this review. However, one of the intended outcomes of this study will be an overview of process factors related to achievement per country.

Educational effectiveness models do not only focus on cognitive outcomes; non-cognitive outcomes like the attitudes, motivation or self-confidence of students could be considered as output factors as well. However, it can also be argued that these affective factors are student characteristics for which students' achievement need to be controlled for. Self-confidence for

example, could also be regarded as indicator of prior achievement in cases where a true measure of prior achievement is unavailable, which is the case in TIMSS (Dumay & Dupriez, 2007; Kaya, et al., 2009; Xin, Xu, & Tatsuoka, 2004). Up till now, the cognitive outcome achievement in core subjects has dominated educational effectiveness research (Scheerens, et al., 2007). In this review, only articles in which the TIMSS-test is the output factor, will be included. Studies which focus only on the attitudes of students as dependent or independent variables, or only on other student background characteristics, input and/or context factors are excluded from this systematic review. The factors mentioned in Table 1 and 2 will be the starting point for the identification and categorization of the factors found based on the secondary analyses of the TIMSS-data sets.

4. Method

4.1 Selection procedure

To gather references to published scientific articles on secondary analyses of TIMSS data, the following online databases were used: Web of science (www.isiknowledge.com), Scopus (www.scopus.com), Picarta (a Dutch electronic database, www.picarta.nl), ERIC and Psycinfo (provided through Ebscohost). All databases were searched using the following search key: “timss OR "trends in international mathematics and science" OR "third international mathematics and science". In all databases is searched in the title, the abstract or the keywords. The search was carried out in February 2010. In total, more than 1600 hits were found (see Table 3). After removing the duplicate articles 988 unique articles were identified.

Table 3 Results literature search (February 2010)

Database	**Number of hits**
ERIC	636
PiCarta	385
PsycInfo	181
Scopus	274
Web of Science	168
Total	**1664**
Duplicates	656
Total number of possible relevant articles	**988**

Five inclusion criteria were formulated to select the studies for the review. To evaluate the usefulness of these criteria a scoping review was done on 25 studies. This evaluation determined the order in which the criteria were used. The following five criteria have been used:

1. The study must be published in peer reviewed journal.
   Non research articles like book reviews and editorials or popular articles (almost half of all references) were excluded at this criterion.
2. The language of the study must be in English.
3. The analyses must be done on 'regular' TIMSS-data.
   Data from the TIMSS-video study and the performance test (1995) were excluded at this criterion. It turned out that in many studies (almost 40%) in which 'TIMSS' was referred to, the data itself were not used for analyses.
4. The dependent variable of the study must be the mathematics or science achievement of the students. Studies in which only descriptive analyses were done or studies in which the dependent variable was not achievement but other variables such as the attitudes of students, were excluded.
5. Of all the factors included in the study, at least one is a manipulative process factor on classroom or school level. Articles which focused only on input, context or student characteristics (or combinations of these factors) were excluded. This criterion resulted in a further reduction of the selection by more than a half.

Bibliographic information like abstract, title, journal title, etcetera was used for the evaluation of the articles based on the first four criteria. When it was not clear whether an article full-filled a certain criterion, the article continued to be part of the selection. For most of the remaining articles it was necessary to review the full text of the article to find out if it also fulfilled the fifth criterion. Two reviewers evaluated each article for the final criterion. In

Table 4 an overview of the selected articles of each step is given. From the 988 articles found in this search, only 61 satisfied the inclusion criteria. That is 6% of all articles.

Table 4 Overview of the selected articles with each criterion

Inclusion criteria	**N**
Peer reviewed articles	514
Language English	442
Based on TIMSS data	271
Dependent variable mathematics or science Achievement	164
At least one process factor on school or class level	61

4.2 Analyzing the data

For each of the 61 studies selected, a data-extraction form was filled in. The data-extraction form had two goals:

1) Categorising the remaining studies in year of data collection, subject, population (age and country), number of citations and journal impact factor (research question 1 and 2)
2) Assessing the quality of the remaining studies, using the quality criteria for quantitative studies as are formulated by Petticrew and Roberts (2006, research question 2). Studies with a 'low' quality were excluded from the final review.

The articles were evaluated by three reviewers. A part of the articles were assessed by two reviewers to check for inter rater reliability. To evaluate the quality of the study two criteria played a decisive role in the final judgement. In accordance with Scheerens et al. (2007) achievement measures should be controlled for student characteristics. Secondly, the study should take into account the specific research design of TIMSS; the so called nested design. In TIMSS, the schools are sampled and within each school one class (in most countries). Therefore, students are not randomly sampled but are nested in classes. This means that a study either should use a method like hierarchical linear modelling (multilevel analyses) or adjust the standard errors for clustering effects (f.e. the jack-knife procedure). However, it must be noted that pad-analyses in a multilevel structure (for analysing direct and indirect effects), were not possible till recently (Muthén & Muthén, 1998-2010). However, if a study conducted pad-analyses (PLS or LISREL) without mentioning the clustering effects it was still considered of low quality. Because of the large number of potentially influencing process factors, this review study will limit itself to only direct effects on achievement. Also, interaction effects were not included in the final analyses.

In summary: if the nested design of TIMSS was not taken into account and/or there was no use of control variables such as student background characteristics (e.g. SES) or context variables (e.g. average SES of the school) without argumentation or reflection, the articles were categorized as *low* quality.

In general, an article was considered *sufficient* if it met the criteria above but had **one** of the following quality issues:

- Only the effect of individual items is analysed in cases where composites or indexes seem to be possible (such as items referring to liking math) without argumentation.
- The level of analyses is different from the level of conclusion/implications. For example, if analyses on student level resulted in conclusions on class (instructional) level.
- It is not clear how the different factors were operationalised and measured in TIMSS (including the TIMSS-test itself). It turned out that the labelling of the same factors could differentiate between authors, or that the same factors were not operationalised in the same way. This means that the author should be clear which variables (composites and items) were included in the analyses.
- There was no information about the number of students/schools included in the analyses.

If more than one of these or other quality issues were found, the article was also indicated as "low quality".

If a study was considered sufficient or high quality, the factors studied in the article were categorized based on the factors, Scheerens et al. (2007) distinguished (research question 3). When appropriate, a factor was added as a new category in the data-extraction form. The effects of the factors were also described (positive, negative, level of significance).

The next section starts with a description of the characteristics of the studies that based on TIMSS-data analysed the relationship of manipulative process factors (either on the classroom or school level) with achievement. This means all 61 studies. The second part of this section presents the results of the quality assessment; the third part describes the conceptual analysis of the analysed factors. For this last part only the studies that were considered to be sufficient or high quality were included.

5. Preliminary results

5.1 Characteristics of the sample

In Figure 1 an overview is given of the number of the selected articles per year. The most productive year was 2008. In this year a special issue was published regarding secondary analyses of the TIMSS-data in the journal “Studies in Educational Evaluation”. Also in 2005 a special issue was published in “Educational Research and Evaluation”. After 2003 the number of TIMSS articles regarding secondary analysis on process factors seems to be increasing. However, in 2009 the number declined again. In general, the TIMSS data seems to be used more and more for secondary analysis related to educational effectiveness.

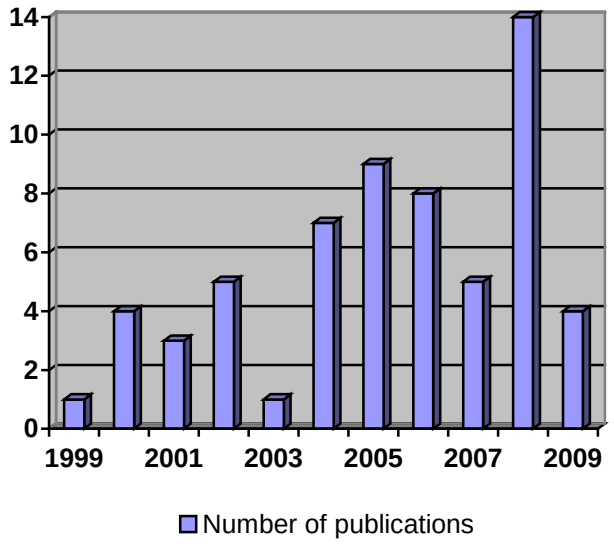

Figure 1 *Distribution of articles in our sample over the years 1996–2010 (n = 61)*

Authorship, citations and journals

Seventy-four unique authors were identified in the sample of 61 articles. The most productive author in this sample has written 14 articles (J. D. House). A majority of the authors only have contributed to one article regarding the topic of educational effectiveness in relation to TIMSS data. Ten of the 74 authors have contributed to more than one article in this sample. This does not mean that these authors could not have written TIMSS-related articles on other subjects, like the attitudes of students. However, generally there seem to be only a few authors who systematically use the TIMSS data for secondary analysis in relation to school and classroom process factors.

Based on the data of the citation index of the Web of Science it was determined (May 2010) how often the 61 selected articles were cited in other peer reviewed articles. The number of citations in Scopus or Google scholar has not been included in these numbers. Therefore, the actual amount of citations is probably higher because both Scopus and Google Scholar take into account the citations from others sources than the articles from the ISI-journals. The 61 articles together were cited 108 times. On average, an article in this sample has been cited 1.77 times, 35 of the 61 articles have not been cited. Four of the 61 articles were cited more than 5 times. The article of Baker from 2002 has been cited the most times, 24 times. As mentioned before, this study only describes the characteristics of a selected set of articles which focus on the influence of process factors on the classroom or school level. Numbers concerning the citations of TIMSS-articles are different if also articles are included which focus on input or context characteristics. A good example is the article of Baker, Akiba, LeTrendre and Wiseman (2001) which focuses on the influence of shadow education and is cited 18 times. However, the impact of the articles that use TIMSS data to study process factors that influence achievement seems to be limited. Only few articles are cited frequently. The majority of the articles have not been cited at all (yet).

The 61 articles were published in 31 journals. Sixteen of these 37 journals are ISI-journals and have a so called impact factor. An impact factor reflects the average number of citations from articles published in an ISI-journal. In Table 5 an overview is given of the journals which published more than one TIMSS article related to process factors.

Table 5 Overview of the journals which published more than one selected TIMSS article, the number of citations and impact factor

Journal name	Total articles	Total citations*	Impact factor
Education	2	0	-
Educational Research and Evaluation	9	9	-
International Journal of Educational Research	2	0	-
International journal of instructional media	8	7	-
Journal of Educational Research	2	18	0.667
Prospects: Quarterly Review of Comparative Education	4	1	-
Studies in Educational Evaluation	10	12	-

Note: * based on citation figures of the Web of Science.

5.2 Characteristics of the TIMSS-data used

Tables 6a, 6b and 6c describe some of the characteristics of the TIMSS data which were used in the 61 studies. A majority of the studies use data from TIMSS-1995 or TIMSS-1999. Only three of the articles used data of more than one TIMSS-study. Grade 8 students are mostly studied in the selected articles. This can be partly explained by the fact that the population of TIMSS-1999 was only grade 8 students. However, in general, there seems to be little attention for grade 4 students as only nine studies focused on this age group (Table 6b). There is also little attention for the possible relation between classroom or school process factors and science achievement. The majority of the selected articles used the TIMSS mathematics achievement scores as output measure (Table 6c).

Table 6a Overview of the TIMSS data used in the articles

TIMSS-study	Number of studies
TIMSS-1995	22
TIMSS-1999	27
TIMSS-2003	14
TIMSS-2007	1

Table 6b Overview of the research population used

Population	Number of studies
4 (3/4)	9
8 (7/8)	52
12 (pre-university or final year)	2

Table 6c Achievement scores studied in articles

Achievement scores	Number of studies
Mathematics	37
Science	14
Both achievement scores	10

Besides the characteristics mentioned above, the number of countries included in a study was also determined. Twenty-one of the 61 studies used the data of more than one country in their analyses. Countries which were often studied in these articles were Japan, the United States and Australia.

5.3 Assessment of the quality

Based on the evaluation of the quality of the studies 26 were selected for the conceptual analyses (see section 4.2). Three articles were excluded from this set, because they didn't comply with the inclusion criteria after further review. The remaining articles were qualified as *low* quality. In 50% of these articles the authors did not take the nested design of TIMSS into account (and provided no argumentation) and in 47% of these articles the authors did not include any control variables and provided no argumentation for not doing so.

In 92% of the articles that were assessed as *sufficient* (13 articles), a type of multilevel analyses were applied. All articles that were considered to be of *high* quality used some form of hierarchical modelling (in most cases HLM of MLwin). Of all 61 studies; 21% were labelled 'high quality'.

In the articles under review, a distinction can be made between three different EER approaches: the economic approach, the psychological/sociological approach and the generalist-educationalist approach (Kyriakides, 2006). In Table 7 an overview is given of the 61 articles by the result of the quality assessment and by approach of EER.

Table 7 Results of the quality assessment of the selected articles on process factors, by approach (n=61)

	Quality		
Approach	*low*	*sufficient*	*high*
Economic	3	1	1
Sociological*	28	0	0
Generalist-educationalist	1	12	12
Total	32	13	13

The economic approach is characterized by the use of so called 'education production models', in which it is assumed that increased inputs will lead to increased outputs. Often, in this approach, there is no distinction made between manipulative and non-manipulative (control or contextual) input factors. Another characteristic is that often no composites or indexes were used, instead the data was analysed on item level. In total, five of the reviewed articles were written from this economic perspective. Only two of these articles used a multilevel education production model.

The second approach focuses on factors at the student level; and includes a further distinction between psychological researchers (focus mainly on affective variables) and sociological researches who focus mainly on student background factors (SES or gender) and on process factors that emerged from the organizational theories (such as school climate or school contextual characteristics). The main characteristic of this approach is that no multilevel structure is used. This does not mean that articles written from this perspective were automatically categorized as low quality. Often, the nested design of TIMSS was taken into account by using the jack-knife procedure. In this selection, none of these articles were indicated as high quality, due to other quality issues. None of the selected articles could be characterized as 'psychological'.

In the generalist-educationalist approach findings of school effectiveness research, teacher effectiveness research and early input-output studies are combined, resulting in models with a multilevel structure (Kyriakides, 2006). Of all the articles written with this perspective in mind, only one was indicated as low quality (Table 7).

5.4 Conceptual analyses

The data extraction forms were used as a starting point for the conceptual analyses. Next to a table with questions about the quality, these forms include an overview of all process factors

mentioned by Scheerens (2007). For this, Table 1 (process factors on school level) and Table 2 (process factors on teacher level, see section 3) were combined.

Tables 8a (grade 8), 8b (grade 4) and 8c (grade 12) show the results of the analyses for studies of sufficient and high quality. Only indicators that showed a significant effect (<.05) in at least one study, country or data set were included, also if the indicator showed no significant effect in another study, country or data set. An overview per article of all indicators of process factors that were included is presented in Appendix A. The direction of the effects is also indicated in the tables below. If the results were contradicting (in one country, subject or study a positive significant effect, in another a negative significant effect) the indicators are presented in bold. In some studies, the answering categories of variables were analysed as dummies. Often, these variables showed no linear effect. For example, no effect for 'not' working in small groups, a positive effect for 'sometimes' working in small groups and a negative effect for 'frequently' working in small groups. In the tables, these indicators are underlined.

Table 8a Results of the systematic review of TIMSS-studies on process factors derived from the school effectiveness model of Scheerens (2007), grade 7 and/ or 8

Process school level	***indicator with a significant effect (< .05)***
Achievement orientation and expectations	*Periods scheduled for teacher to counsel students*
Educational Leadership	Principal responsible for community relationships
Consensus and cohesion among staff	Staff-cooperation Teachers use of professionalization activities *Teachers meet regularly to discuss instructional goals and issues*
Curriculum quality/ Opportunity to learn	Number of topics taught School mean teachers emphasis on homework Students tracked in top stream *Students tracked in bottom stream* **No streaming or tracking**
School climate	Academic climate Safety at school (teachers' perception) Cheating of students *Absence of students* *Violations of dress code* *Students causing injury to other students* *Frequency school has to deal with class disturbance* *Problems with late arrival on school*
Evaluative potential	
Parental involvement	Parental involvement (all items) Direct parental involvement School expects parents to ensure child's homework completion
Structured instruction	
Differentiation, adaptive instruction	School mean relating to everyday life School mean trying to solve an example related to a new topic School mean having the teacher ask us what we know related to new topic
Feedback and reinforcement	

Table 8a (continued) Results of the systematic review of TIMSS-studies on process factors derived from the school effectiveness model of Scheerens (2007), grade 7 and/ or 8

Process class level	***indicator with a significant effect (< .05)***
Achievement orientation and expectations	
Curriculum quality/ Opportunity to learn	Amount of homework Providing practice and application opportunities Using various teaching means Minutes of instruction per week in subject Use of inquiry in science Teacher's emphasis on understanding concepts Using calculator for solving problems Using calculator for explorations Using calculator for checking Using calculator for test Using calculator for routine procedures *Extent of calculator use* *Percentage no homework in class* **Amount of time using textbook**
Class climate (creation and dimensions of)	Class climate Class mean attitude toward math Class mean self-confidence *Disruptive students limit teaching in maths class* *Perceived limitations in teaching due to problem students* *Class mean uncontrollable attribution* **Class mean self pressure**
Evaluative potential	*Having quiz or test*
Effective learning time	*Student self rated attentiveness* *Lesson interruption*
Structured instruction	Practicing of basic mathematics operations without calculator Whole class teaching Listening to the teacher *How often students are asked to provide explanations* Looking at the textbook while the teacher talks about it Having the teacher explain the rules and definitions
Differentiation, adaptive instruction	Providing opportunities for collaboration *Working out problems on their own* **Working in small groups** **Relating to everyday life** Having the teacher ask us what we know related to the new topic
Feedback and reinforcement	Teacher's time spent on scrutiny of exams/tests Discussion on homework made

Note: Normal= positive effect in at least one country or one subject, without contradicting negative effects
Italic= negative effect in at least one country or one subject, without contradicting positive effects
Bold= contradicting results for countries, subjects or TIMSS-year (positive and negative effects)
Underlined=effect is not linear

Table 8b Results of the systematic review of TIMSS-studies on process factors derived from the school effectiveness model of Scheerens (2007), grade 4

Process school level	***indicator with a significant effect (< .05)***
Achievement orientation and expectations	
Educational Leadership	
Consensus and cohesion among staff	
Curriculum quality/ Opportunity to learn	Number of topics taught Ability grouping
School climate	
Evaluative potential	
Parental involvement	
Structured instruction	
Differentiation, adaptive instruction	
Feedback and reinforcement	
Process class level	
Achievement orientation and expectations	
Curriculum quality/ Opportunity to learn	**Use of inquiry in science**
Class climate (creation and dimensions of)	Teachers' support perceived by students Class mean self–confidence
Evaluative potential	
Effective learning time	
Structured instruction	
Differentiation, adaptive instruction	*Student oriented teaching*
Feedback and reinforcement	

Note: Normal= positive effect in at least one country or one subject, without contradicting negative effects
Italic= negative effect in at least one country or one subject, without contradicting positive effects
Bold= contradicting results for countries, subjects or TIMSS-year (positive and negative effects)

Table 8c. Results of the systematic review of TIMSS-studies on process factors derived from the school effectiveness model of Scheerens (2007), grade 12

	indicator with a significant effect (< .05)
Process school level	
Achievement orientation and expectations	
Educational Leadership	
Consensus and cohesion among staff	
Curriculum quality/ Opportunity to learn	
School climate	*Schools' report on incidents with students*
Evaluative potential	
Parental involvement	
Structured instruction	
Differentiation, adaptive instruction	
Feedback and reinforcement	
Process class level	
Achievement orientation and expectations	
Curriculum quality/ Opportunity to learn	
Class climate (creation and dimensions of)	
Evaluative potential	
Effective learning time	
Structured instruction	*Passive teaching*
Differentiation, adaptive instruction	Active teaching
Feedback and reinforcement	

Note: Normal= positive effect in at least one country or one subject, without contradicting negative effects
Italic= negative effect in at least one country or one subject, without contradicting positive effects
Bold= contradicting results for countries, subjects or TIMSS-year (positive and negative effects)

Table 8a,b,c show that indicators of many factors that are distinguished by Scheerens (2007) are also studied in the selected TIMSS-studies. In the TIMSS Curriculum Framework, on which both test and questionnaires are based; it is assumed that the attained curriculum (student achievement) is the result of the intended (what should be taught) and implemented curriculum (what is actually taught) and the result of country, school, class and student level characteristics. Because of this curriculum framework, a lot of attention is paid to collect information on curriculum related indicators. Therefore, relatively a lot of studies used indicators of curriculum quality' and 'opportunity to learn' on both school and class level. Especially opportunity to learn (often operationalised as number of topics taught) turned out to be an important positive factor in most countries (although not in all countries, see also Table 9).

Table 8 also shows that class and climate factors are often included in studies. Composites or indexes as well as individual items are used to study the effect of class and school climate on achievement. Table 8a,b,c as well as Appendix A show that most of the effects found are comparable across the different studies. A less clear effect is the class mean of experienced self pressure (O'Dwyer, 2005). In some countries a positive effect was found, in other countries a negative effect. Other school- or class climate factors show that an orderly, safe school or classroom environment with few disruptive incidents has a positive effect on achievement. A notable exception is the positive effect (where a negative effect would be expected) found in the study of Pugh and Telhaj (2008) regarding the cheating of students.

Process level factors related to instruction (structured instruction, differentiated, adaptive instruction) are relatively often included in the selected studies as well. Most articles used the individual items opposed to composites or indexes to study these factors. It seems that making composites of indicators of instruction measured in TIMSS is more difficult compared to, for example, school climate. Furthermore, the effects of a number of these indicators turned out to be contradicting or not linear. Table 8a shows that the effects found for these factors (for instance working in small groups) is different across the different studies or even within a study. Both negative and positive effects are found. The study of Ma and Papanastasiou (2006) indicate that possibly the intensity of these activities determine the positive or negative influence on achievement.

Finally, there seems to be little attention for a number of process factors on both school and classroom level. These factors are: achievement orientation and expectations, educational leadership, evaluative potential and reinforcement and feedback. Until now, these factors are not or limitedly operationalised in the TIMSS-questionnaires.

Table 9 gives an overview of the effects of the most researched process factors per country. A positive sign is given to a country when at least one positive effect has been found and no negative effects have been found.

Table 9 *Effects of selected process factors per country*

	Classroom level factors							School level factors		
Countries	Amount of homework	Perceived limitations in teaching due to problem students	Class mean self pressure	Class mean attitude toward math	Teachers's support perceived by students	Whole class teaching	Working together in small groups	Number of topics taught	Academic climate	Safety at school (teachers' perception)
Australia	+		+	+	o	o	o	+		
Bulgaria								+		
Canada	o		+	o		o	+/-	+		
Chile								+		
Cyprus	o		o	o						
Czech Rebublic	+		-	+						
Estonia										
England		-						+	o	+
Flemish Belgium	+	-	-	+				+	+	+
France						o	o	o		
Germany						o	o	o		
Hong Kong	+		+	+		+	-	o		
Hungary	+		o	o						
Iceland						o	o	o		
Iran	+		-	o						
Irsael	o		o	o			-			
Japan					o					
Korea	o		o	+		o	o	o		
Latvia	+		o	+						
Lithuania	o									
Netherlands	o	o	-	+						
New-Zealand	o		o	o				+	o	+
Slovakia	o		-	+						
Slovenia	o		o	o						
Romania	+		o	+						
Russia	o		o	+					+	
Scotland					o					
Singapore	o				+	o	o			
South-Africa			o	+				+		
South-Korea										
Thailand	+		+							
U.S.	+		o	+	+	o	o	+	o	o

Note: o = no effect found
+ = positive effect
- = negative effect

Tables 8a,b and c showed that the highest number of the analysed TIMSS-variables are categorised as indicators of curriculum quality and opportunity to learn on both school and class level. Although there are a number of studies or countries were no effect was found for the number of topics taught (five countries, see Table 9), this indicator of 'opportunity to learn' on school level seem to have a positive influence on achievement. For eight countries a positive effect was found.

The amount of homework given by the teacher is an indicator for opportunity to learn on classroom level. This factor also has been analysed in the selected studies for at least 21 countries. The positive effect of this factor is less clear than the number of topics taught at the school level. The results are mixed; in 10 countries a positive effect was found, in 11 countries no effect was found. Furthermore, the studies selected have not found positive or negative effects for the type of homework given. One explanation could be that in some countries (such as the Netherlands in grade 4) homework is assigned to students who are falling behind. It seems that more attention should be given to conditions in which giving homework has a positive influence of on achievement.

6. Conclusions

The aim of this study was to give an overview of studies based on TIMSS data which contribute to theories of Educational Effectiveness Research (EER). This was done by systematically analyzing these studies as they are represented in scientific journals since the release of the TIMSS-1995 data. The method of this systematic review was based on that of Petticrew and Roberts (2006) and included a selection of relevant articles based on the framework of school effectiveness and an assessment of the quality. The final selection (studies which were indicated as sufficient or high quality) was analysed further to find out to what extent the results of these studies supported the list of Scheerens et al.(2007) of effective process factors on school and class level.

The selection procedure started with almost a thousand articles, of which only 61 were selected because they met all selection criteria. Based on the evaluation of the quality of the studies, only 26 articles were selected for the conceptual analyses. Of this final selection, two studies were conducted from an economic approach (Kyriakides, 2006). The remaining articles could be characterized as the generalist-educationalist approach. In most of the studies that were considered to be of high quality, multilevel analyses were applied.

There are a number of remarks than can be made about the studies in this systematic review. First, over the years, the TIMSS data seems to be used more and more for secondary analysis related to educational effectiveness. The number of studies in which TIMSS-data is used (271 studies), but the achievement test is **not** the dependent variable (107 studies), has become smaller than the number of studies in which achievement is the dependent variable (164 studies). This seems to be a more positive result compared to the conclusion of Beaton and Robitaille in 2002 . However, of these 164 studies, less than 40% included indicators of process factors.

Second, it turned out that most of the 61 TIMSS-studies used grade 7 and/or grade 8 data. The low number of grade 12 or final year studies was to be expected, because the grade 12 study was only conducted in 1995 and 2008 (TIMSS-Advanced 2008) and the final year study in 1995. However, although grade 4 was not part of TIMSS-1999, grade 4 students were assessed in 1995, 2003 and 2007. It seems that researchers using TIMSS for secondary analyses on factors related to achievement, are more interested in lower secondary education

than primary education. The same goes for mathematics compared to science. The majority of the selected articles used the TIMSS mathematics achievement score as output measure. In the final selection (after the quality assessment), only one study on science achievement of grade 4 students remained (Kaya, et al., 2009). In terms of the contribution of TIMSS to education effectiveness research, the contribution seems to be limited to lower secondary education. Furthermore, the impact of these studies seems to be limited as well. Only a few articles were cited frequently while the majority of the 61 articles have not been cited at all (yet).

Most of the selected 61 studies used the data of European or Asian countries. In most case these are not the countries that performed extremely low in the international ranking lists of TIMSS. One would expect that countries with very low performance would receive more attention of researchers in order to find ways of improving the educational level. There were a few studies, all by the same author, using the data of the lowest scoring country of TIMSS-2003, South Africa (Howie, 2004; Howie, 2005a; Howie, 2005b; Howie & Scherman, 2008; Howie, Scherman, & Venter, 2008). From the perspective of the school effectiveness model, analyses of TIMSS-data of Arabic or African countries would be an interesting addition because this model is based on developed countries as well. Furthermore, the focus on European and Asian countries does not mirror the participation of countries in TIMSS.

During the review process, one of the quality criteria became more important than expected beforehand: the author should be clear which variables (composites and items) were included in the analyses. It turned out that the labelling of the same factors quite often differentiated between authors, or that the same factors were not operationalised in the same way. For example, the indicator ‘number of topics taught’ was labelled as opportunity to learn, number of topics taught, or content coverage. In several cases the labelling and operationalisation of variables was not clear and as a consequence the results could not or only partly used in the conceptual analyses.

Many of the effective enhancing process factors that are distinguished by Scheerens (2007) can also be found in the selected TIMSS-studies. Most results were also in line with the results of the meta analyses of Scheerens et al.(2007). However, a number of process factors received no or little attention. Three of these factors, reinforcement and feedback, educational leadership and evaluative potential are included in the Contextual framework of TIMSS-2011

(Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009). Recently, the evaluative potential receives an increasing interest from the field of EER (Schildkamp, Visscher, & Luyten, 2009; Verhaeghe, Vanhoof, Valcke, & Van Petegem, 2010).

Finally, the systematic review described in this paper is not complete. The next step in this research will be the inclusion of peer reviewed TIMSS-articles in books and dissertations. Furthermore it must be noted that the use of only peer reviewed articles could result in publication bias. Studies in which no significant effects are found, or are not able to explain a certain percentage of variance, are less likely to be published. No effects could be relevant as well, especially as this study has shown that the relation between manipulative factors and achievement could differ a lot between countries or educational systems.

References

Ammermuller, A., Heijke, H., & Wossmann, L. (2005). Schooling quality in eastern europe: Educational production during transition. *Economics of Education Review, 24*(5), 579-599. doi: 10.1016/j.econduurev.2004.08.010

Baker, D. P., Akiba, M., LeTendre, G. K., & Wiseman, A. W. (2001). Worldwide shadow education: Outside-school learning, institutional quality of schooling, and cross-national mathematics achievement. [Article]. *Educational Evaluation and Policy Analysis, 23*(1), 1-17.

Baker, D. P., Goesling, B., & Letendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-loxley effect" On mathematics and science achievement. *Comparative Education Review, 46*(3), 291-312.

Bankov, K., Mikova, D., & Smith, T. M. (2006). Assessing between-school variation in educational resources and mathematics and science achievement in bulgaria. *Prospects: Quarterly Review of Comparative Education, 36*(4), 447-473. doi: 10.1007/s11125-006-9005-7

Beaton, A., & Robitaille, D. (2002). A look back at timss: What have we learned about international studies? In D. Robitaille & A. Beaton (Eds.), *Secondary analysis of timss-data* (pp. 409-417). Dordrecht: Kluwer.

Bos, K. T. (2002). *Benefits and limitations of large-scale international comparative achievement studies : The case of iea's timss study.* University of Twente, Enschede.

Diepen, M. van. (2007). *Reading literacy development from an international perspective* Radboud University, Nijmegen.

Dumay, X., & Dupriez, V. (2007). Accounting for class effect using the timss 2003 eighth-grade database: Net effect of group composition, net effect of class process, and joint effect. *School Effectiveness and School Improvement, 18*(4), 383-408. doi: 10.1080/09243450601146371

Howie, S. (2004). A national assessment in mathematics within an international comparative assessment. *Perspectives in Education, 22*(2), 149-162.

Howie, S. (2005a). Contextual factors at the school and classroom level related to pupils' performance in mathematics in south africa. *Educational Research and Evaluation, 11*(2), 123-140. doi: 10.1080/13803610500110703

Howie, S. (2005b). System-level evaluation: Language and other background factors affecting mathematics achievement. *Prospects: Quarterly Review of Comparative Education, 35*(2), 175-186. doi: 10.1007/s11125-005-1820-8

Howie, S., & Scherman, V. (2008). The achievement gap between science classrooms and historic inequalities. *Studies in Educational Evaluation, 34*(2), 118-130. doi: 10.1016/j.stueduc.2008.04.007

Howie, S., Scherman, V., & Venter, E. (2008). The gap between advantaged and disadvantaged students in science achievement in south african secondary schools. *Educational Research and Evaluation, 14*(1), 29-46. doi: 10.1080/13803610801896380

Howie, S. J. (2002). *English language proficiency and contextual factors influencing mathematics achievement of secondary school pupils in south africa.* University of Twente, Enschede.

Istrate, O., Noveanu, G., & Smith, T. M. (2006). Exploring sources of variation in romanian science achievement. *Prospects: Quarterly Review of Comparative Education, 36*(4), 475-496. doi: 10.1007/s11125-006-9006-6

Jurges, H., Buchel, F., & Schneider, K. (2005). The effect of central exit examinations on student achievement: Quasi-experimental evidence from timss germany. *Journal of the European Economic Association, 3*(5), 1134-1155. doi: 10.1162/1542476054729400

Jurges, H., & Schneider, K. (2004). International differences in student achievement: An economic perspective. *German Economic Review, 5*(3), 357-380.

Kaya, S., & Rice, D. C. (2009). Multilevel effects of student and classroom factors on elementary science achievement in five countries. *International Journal of Science Education*. doi: 10.1080/09500690903049785

Kyriakides, L. (2006). Using international comparative studies to develop the theoretical framework of educational effectiveness research: A secondary analysis of timss 1999 data. *Educational Research and Evaluation, 12*(6), 513-534. doi: 10.1080/13803610600873986

Lamb, S., & Fullarton, S. (2002). Classroom and school factors affecting mathematics achievement: A comparative study of australia and the united states using timss. *Australian Journal of Education, 46*(2), 154-171.

Ma, X. (2001). Stability of socio-economic gaps in mathematics and science achievement among canadian schools. *Canadian Journal of Education, 26*(1), 97-118. Retrieved from http://www.csse.ca/CJE/Articles/FullText/CJE26-1/CJE26-1-Ma.pdf

Ma, X., & Papanastasiou, C. (2006). How to begin a new topic in mathematics: Does it matter to students' performance in mathematics? *Evaluation Review, 30*(4), 451-480. doi: 10.1177/0193841X05284090

Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., & Shen, C. (2000). Effective schools in science and mathematics *IEA's Third International Mathematics and Science Study* Retrieved from http://timss.bc.edu/timss1995i/TIMSSPDF/T95_EffSchool.pdf

Meelissen, M., & Luyten, H. (2008). The dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation, 34*(2), 82-93. doi: 10.1016/j.stueduc.2008.04.004

Mere, K., Reiska, P., & Smith, T. M. (2006). Impact of ses on estonian students' science achievement across different cognitive domains. *Prospects: Quarterly Review of Comparative Education, 36*(4), 497-516. doi: 10.1007/s11125-006-9007-5

Mullis, I., Martin, M., & Foy, P. (2008). *Timss 2007 international mathematics report: Findings from iea's trends in international mathematics and science study at the fourth and eighth grades*: TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *Timss-2011 assessment frameworks*. Boston: TIMSS & PIRLS International Study Center, Lynch school of education, Boston college.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

O'Dwyer, L. (2005). Examining the variability of mathematics performance and its correlates using data from timss '95 and timss '99. *Educational research and evaluation : an international journal on theory and practice, 11*(2), 155-178. doi: 10.1080/13803610500110802

Papanastasiou, C. (2008). A residual analysis of effective schools and effective teaching in mathematics. *Studies in Educational Evaluation, 34*(1), 24-30. doi: 10.1016/j.stueduc.2008.01.005

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide* Oxford: Blackwell.

Pong, S.-l., & Pallas, A. (2001). Class size and eighth-grade math achievement in the united states and abroad. *Educational Evaluation and Policy Analysis, 23*(3), 251-273. doi: 10.3102/01623737023003251

Postlethwaite, T. N. (1995). International empirical research in comparative education: An example of the studies of the international association for the evaluation of educational achievement (iea). *Journal für Internationale Bildungsforschung, 1*(1), 1-19. Retrieved from http://213.198.68.42/zs/tc1-95/postleth.pdf

Postlethwaite, T. N., & Ross, K. N. (1992). Effective schools in reading: Implications for educational planners. An exploratory study. The Hague: International Association for the Evaluation of Educational Achievement.

Pugh, G., & Telhaj, S. (2008). Faith schools, social capital and academic attainment: Evidence from timss-r mathematics scores in flemish secondary schools. *British Educational Research Journal, 34*(2), 235-267. doi: 10.1080/01411920701532178

Ramirez, M.-J. (2006). Understanding the low mathematics achievement of chilean students: A cross-national analysis using timss data. *International Journal of Educational Research, 45*(3), 102-116. doi: 10.1016/j.ijer.2006.11.005

Rodriguez, M. C. (2004). The role of classroom assessment in student performance on timss. *Applied Measurement in Education, 17*(1), 1-24. doi: 10.1207/s15324818ame1701_1

Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science Ltd.

Scheerens, J., & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13*(7), 691-706. doi: 10.1016/0883-0355(89)90022-0

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: University of Twente, Department of Educational Organisation and Management.

Scheerens, J., Seidel, T., Witziers, B., Hendriks, M., & Doornekamp, G. (2005). *Positioning the supervision frameworks for primary and secondary education of the dutch educational inspectorate in current educational discourse and validating core indicators against the knowledge base of educational effectiveness research*. Enschede/Kiel: University of Twente/Institute for Science Education (IPN)

Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement, 20*(1), 69-88. doi: 10.1080/09243450802605506

Schreiber, J. B. (2002). Institutional and student factors and their influence on advanced mathematics achievement. *Journal of Educational Research, 95*(5), 274-286. doi: 10.1080/00220670209596601

Schreiber, J. B., & Chambers, E. A. (2003). American high school seniors' mathematics literacy achievement. *North American Journal of Psychology, 5*(1), 15-30. Retrieved from EBSCOhost Psychology and Behavioral Sciences Collection

Tarr, J. E., Mittag, K. C., Uekawa, K., & Lennex, L. (2000). A comparison of calculator use in eighth-grade mathematics classrooms in the united states, japan, and portugal: Results from the third international mathematics and science study. *School science and mathematics : official journal of the School Science and Mathematics Association, 100*(3), 139-150.

Van den Broeck, A., Opdenakker, M.-C., & Van Damme, J. (2005). The effects of student characteristics on mathematics achievement in flemish timss 1999 data. *Educational Research and Evaluation, 11*(2), 107-121. doi: 10.1080/13803610500110745

Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010). Using school performance feedback: Perceptions of primary school principals. *School Effectiveness and School Improvement, 21*(2), 167-188. doi: 10.1080/09243450903396005

Xin, T., Xu, Z., & Tatsuoka, K. (2004). Linkage between teacher quality, student achievement,and cognitive skills: A rule-space model. *Studies in Educational Evaluation, 30*(4), 205-223. doi: 10.1016/j.stueduc.2004.09.002

Zhao, H., & Akiba, M. (2009). School expectations for parental involvement and student mathematics achievement: A comparative study of middle schools in the us and south korea. *Compare, 39*(3), 411-428. doi: 10.1080/03057920701603347

Zuzovsky, R. (2008a). Capturing the dynamics behind the narrowing achievement gap between hebrew-speaking and arabic-speaking schools in israel: Findings from timss 1999 and 2003. *Educational Research and Evaluation, 14*(1), 47-71. doi: 10.1080/13803610801896562

Zuzovsky, R. (2008b). Closing achievement gaps between hebrew-speaking and arabic-speaking students in israel: Findings from timss-2003. *Studies in Educational Evaluation, 34*(2), 105-117. doi: 10.1016/j.stueduc.2008.04.006

Appendix A: Overview of process factors analysed in TIMSS-studies, per article

Author/year	**TIMSS-study**	**Grade**	**Subject**	**Country/Countries**	**Process factors included**
(Ammermuller, Heijke, & Wossmann, 2005)	TIMSS-1995	7/8	Mathematics and Science	Czech Republic Hungary, Slovakia Slovenia, Latvia, Lithuania Romania	Amount of homework
(Baker, Goesling, & Letendre, 2002)	TIMSS-1995	7/8	Mathematics and Science	36 countries countries	*Absence of students* (time principal spent on discussing educational issues with teachers) (percentage of teachers working at the school for more than 5 years)
(Bankov, Mikova, & Smith, 2006)	TIMSS-2003	8	Mathematics + all subjects separately	Bulgaria	Number of topics taught Use of inquiry in science
(Dumay, et al., 2007)	TIMSS-2003	8	Mathematics	Flemish, England, Netherlands, USA	Number of topics taught Academic climate Safety at school (teachers' perception) *Perceived limitations in teaching due to problem students* (emphasis on homework)
(Howie, 2005b)	TIMSS-1999	8	Mathematics	South-Africa	Time teacher spend on lesson planning (minutes of instruction per week)
(Istrate, Noveanu, & Smith, 2006)	TIMSS-2003	8	All science subjects separately	Romania	Academic climate Use of inquiry in science
(Jurges & Schneider, 2004)	TIMSS-1995	7/8	Mathematics	23 OECD-landen	Amount of homework *Perceived limitations in teaching due to problem students*
(Kaya, et al., 2009)	TIMSS-2003	4	Science	Singapore, Japan, USA, Australia, Scotland	Teachers' support perceived by students Class mean self-confidence **Use of inquiry in science** (minutes of instruction per week)
(Kyriakides, 2006)	TIMSS-1999	8	Mathematics	all participating countries	Amount of homework Providing practice and application opportunities

Author/year	TIMSS-study	Grade	Subject	Country/Countries	Process factors included
					Using various teaching means Providing opportunities for collaboration *Student self rated attentiveness* *Lesson interruption* (time spent on private tutoring) (type of homework) (classroom management) (correcting and using completed homework) (using assessment information for formative reasons)
(Lamb & Fullarton, 2002)	TIMSS-1995	7/8	Mathematics	Australia, U.S.	Amount of homework Top stream or track **Amount of time using textbook** **No streaming or tracking** *Bottom stream or track* (% time teaching maths (class)) (% time teaching maths (school)) (behavioural disturbances) (absenteeism) (academically selective) (open admission) (same teacher)
(Ma, 2001)	TIMSS-1995	7/8 3/4	Mathematics Science	Canada	Staff cooperation (disciplinary climate)
(Ma, et al., 2006)	TIMSS-1999	8	Mathematics + all subjects separately	Canada	Having the teacher explain the rules and definitions Looking at the textbook while the teacher talks about it Relating to everyday life Working in small groups Having the teacher ask students what they know related to the new topic Trying to solve an example related to the new topic
(Meelissen, et al., 2008)	TIMSS-2003	4	Mathematics	The Netherlands	Number of topics taught Safety in school (teacher) Ability grouping *Feeling unsafe at school(student)*

Author/year	TIMSS-study	Grade	Subject	Country/Countries	Process factors included
					Student oriented teaching (cognitive safety in class) (class atmosphere)
(Mere, Reiska, & Smith, 2006)	TIMSS-2003	8	Science	Estonia	(science teacher climate) (average inquiry across science teachers)
(O'Dwyer, 2005)	TIMSS-1995 TIMSS-1999	8	Mathematics	23 countries which participated in both TIMSS studies	Amount of homework Class mean attitude toward math **Class mean self pressure**
(Pong & Pallas, 2001)	TIMSS-1995	8	Mathematics	Australia, Canada, France, Germany Hong Kong, Korea Iceland, Singapore, U.S.	Number of topics taught Whole class teaching Discussion about homework *Small group* (individual work)
(Pugh, et al., 2008)	TIMSS-1999	8	Mathematics	Flemish Belgium	Too many individual items for inclusion in table
(Ramirez, 2006)	TIMSS-1999	8	Mathematics	Chile. South Korea, Malaysia, the Slovak Republic and Miami-Dade County	Number of topics taught (curriculum materials and activities developed by school) (subject emphasis (basic and advanced) (additional math course)
(Rodriguez, 2004)	TIMSS-1995	7/8	Mathematics	U.S.	Homework frequency by teacher Class mean self-self-confidence *Percentage no homework in class* *Class mean uncontrollable attribution* (workbook worksheets) (teacher made objective tests)
(Schreiber, 2002)	TIMSS-1995	12	Mathematics	U.S.	Active teaching *Passive teaching* (schools' report on incidents with students)
(Schreiber & Chambers, 2003)	TIMSS-1995	12	Mathematics	U.S.	Active teaching *Schools' report on incidents with students*
(Tarr, Mittag,	TIMSS-1995	8	Mathematics	U.S, Japan,	Using calculator for checking

Author/year	TIMSS-study	Grade	Subject	Country/Countries	Process factors included
Uekawa, & Lennex, 2000)				Portugal	Using calculator for test Using calculator for routine procedures Using calculator for solving problems Using calculator for explorations *Extent of calculator use*
(Van den Broeck, et al., 2005)	TIMSS-1999	8	Mathematics	Flanders-Belgium	(constructivist learning (as perceived by the student)
(Zhao & Akiba, 2009)	TIMSS-1999	8	Mathematics	South-Korea U.S.	Parental involvement (all items) Direct involvement (assisting teachers) (participation in school governance)
(Zuzovsky, 2008a)	TIMSS-1999 TIMSS-2003	8	Mathematics Science	Israel	*Problems with late arrival on school disruptive students behaviors limit teaching* *Working together in small groups* *Working out problems on their own* *Relating to everyday life* *How often students are asked to provide explanations* *Having quiz or test* (reviewing homework) (conducting an experiment)
(Zuzovsky, 2008b)	TIMSS-2003	8	Mathematics Science	Israel	Teachers use of professionalization activities offered by the school Practicing of basic mathematics operations without calculator Listening to the teacher Relating to everyday life

Note: Normal= positive effect in at least one country or one subject, without contradicting negative effects

Italic= negative effect in at least one country or one subject, without contradicting positive effects

Bold= contradicting results for countries of subjects (positive and negative effects)

Underlined=effect is not lineair

Between brackets= no effect at all