

The Study of Science Achievement in Six Countries: Finland, England, Hungary, Japan, Latvia and Russia. Study based on TIMSS–1999

Pasi Reinikainen, Institute for Educational Research, University of Jyväskylä,
pasi.reinikainen@ktl.jyu.fi.

Abstract

This paper introduces Reinikainen's (2007) study which explored country-specific explanatory variables for eighth-grade student science achievement in Finland, England, Hungary, Japan, Latvia and Russia by multi-level modeling of the TIMSS–1999 data. The study was based on a pragmatic approach, in which instead of using one across the countries model or international indicators each country had its own model and explanatory indicators based solely on statistically significant explanatory variables. These variables were presented in graphical form that facilitates interpretation of the results also by persons unfamiliar with multi-level modeling. The study did not only present explanatory variables but also aimed at a new level of interpreting these secondary results: firstly by describing the demographics, educational systems and science education practices of the studied countries, and secondly, by using national science education experts to provide emic interpretation of the statistical results.

Keywords: *science, TIMSS–1999, multilevel modeling, country-specific models, mixed method*

Introduction

There is common agreement among researchers that education, pedagogy, teaching and learning are culturally embedded (e.g. Alexander, 1999; Bempechat et al., 2002; Bos, 2002; Cogan & Schmidt, 1999; Schmidt & McKnight, 1998). The understanding of cultural traditions and specific contexts within different systems is essential in detecting successful educational practices and ideas, and vital if these are planned to be transported or adopted internationally (e.g. Bos, 2002; Kaiser, 1999; Pepin, 2003; Schmidt & Burstein, 1993).

Large-scale International Comparative Achievement Studies (LINCAS) have also been criticized for focusing perhaps too much on producing fragmented international ranking lists of variables (e.g. Kaiser et al., 2002; Keitel & Kilpatrick, 1999; Pepin, 2003; Sjøberg, 2003; Stevenson, 1998) and distinct packages of publications. This can lead to a misrepresentation

of the complexity of the phenomena being researched and of the ways they can be understood (Pepin, 2003). According to Sjøberg (2003) there is also a clear need to complement the valuable data from TIMSS-like studies with more open and culturally sensitive information and perspectives. Also Bos (2002) notice this and wrote that the descriptive ambition of these studies seems to be better met than the understanding function of these studies.

Reinikainen (2007) conducted the literature study on secondary analyses of TIMSS data. It showed at least two things. Firstly, most of the existing further analyses of TIMSS concentrated in mathematics rather than in science. And secondly, both multi- and single country analyses focused largely on theoretical and methodological issues rather than describing and explaining the cultural issues and situations in which these results were actually produced.

However, these national and cultural effects are very strongly present at international studies. For example study of Kyriakides (2006) (three-level HLM) about TIMSS–1999 data showed that 31% of the variance was at the student level, 25% at the teacher level and 44% at the country level. Accordingly to him, this large observed variation at national level might suggest that the educational systems of different countries bear little resemblance, and thus TIMSS data should not be used for summative reasons (i.e. ranking countries according to their student outcomes), but for identifying factors associated with student achievement at various levels.

Also Martin and colleagues (2000b) concluded their Effective schools in science and mathematics study, that it is likely to be more fruitful to use country-level variables in the explanatory models rather than common variables that are expected to operate in the same way across all countries. One of these common variables is “the Positive Attitude Towards Science” (PATS) index, which is perhaps the most commonly used in country-level analyses, and which is also one of these culturally biased variables. The association between PATS and science achievement is positive in positive in within-country analyses, but opposite in cross-country results (e.g. Kyriakides & Caralambous, 2004).

The realization of these previous issues led Reinikainen to use country-specific factors and models to discover something new, unexpected, and perhaps contradictory to previous findings. It also guided him to aim at more descriptive than comparative outcomes. The author also noticed that the cultural distance between the researcher and the participants of the studied countries would decrease the validity of the study and that the study was too extensive to be approached alone. For this reason, Reinikainen chose to use the alternative “group-based” interpretation of mixed methods. Moghaddam and colleagues (2003) wrote that the research enterprise itself is a collective act which generates a view of a given aspect of the world through multitudes of contributions and, thus, the interpretations of the results

should also be performed collectively.

Reinikainen decided to observe a pragmatic worldview in his study, including employing “what works” using diverse approaches and valuing both objective and subjective knowledge. Thus, in this study the research questions were of primary importance. The main research questions in Reinikainen’s study were:

1. What factors explain science achievement in the six countries England, Finland, Hungary, Japan, Latvia and Russia?
2. In what ways do emic interpretations help us to understand the quantitative results?

Methodology

To answer to these questions, Reinikainen used so called Sequential Explanatory Design (Tashakkori & Teddlie, 2003; Creswell & Plano Clark, 2007). The study firstly consisted of two distinct phases: quantitative followed by qualitative, and then the third phase where findings were combined (see Figure 1).

[Insert Figure 1 about here]

The first phase and the emphasis of this study was on country-specific two-level (student = level 1 and school = level 2) HLM models (QUAN) based on the TIMSS–1999 database. These models identified students’ gender, home background, free time, school background, student motivational aspects and affective outcomes and former success related factors, and the directions and strengths in which these were statistically significantly connected with science achievement in the studied countries.

The second phase of the study was the qualitative elaboration of quantitative data. This phase included revision and validation of the HLM models by science education experts from the studied countries and semi-structured interviews (qual) of these experts. Interviews were built on the results of the HLM models with the main purpose of producing emic explanations and interpretations of these predictors of science achievement. Science education experts were chosen who had been previously involved in conducting TIMSS or other LINCAS studies in their native country. These experts were considered not only to be familiar with the TIMSS study, but also, being of some standing in science education in their native countries, to be the best able to provide research based explanations and interpretations of the findings and

possibly able to carry out further research on the findings of this study. The science education experts were also asked to provide information about their school system and, more specifically, about science education in their country.

The results of these two phases were integrated in the reporting stage of the results of the national models (Phase 3). These were presented together with descriptions of countries; their educational systems and science education (see Reinikainen, 2007 for detailed information about the used methodology).

Results

Graphical presentation of HLM-results

As a main result of Reinikainen's study, country specific two-level models of science achievement were created for all six countries. Because of the limited length of this article only the Finnish model is presented graphically (see Figure 2).

[Insert Figure 2 about here]

In existing literature, HLM model results and variable data are typically presented in tabular form. Since not all potential readers will be familiar with HLM models, Reinikainen decided to develop a more illustrative graphical presentation of the results. In this graphical presentation, all of the central information about the model and the explanatory variables is given. It introduces the statistically significant variables in the final and best (the one that explains the largest proportion of total variance in science achievement) country-level models. The presentation shows how different variables are connected with science achievement (positive/negative), how large an impact a certain variable has on the score, and also the classification of variables. What is also useful, and one of the main points of this graphical presentation, is that it also includes the mean values of the explanatory variables along with their minimum and maximum values (endpoints of the bars). In the figures these endpoints are labeled according to the type of student response, and the proportion of students who gave this response is also given. The graphical presentation of HLM results contains a huge amount of data and it can take time some time for the reader to become acquainted with it. However, the results are considerably easier to interpret in graphical form than in tabular form.

Selected national findings and their interpretations

Since the main purpose of this paper is to describe Reinikainen's study, only a few of the

country-level findings from each studied country are selected and presented in this paper. Hopefully readers of this paper will get an idea how emic interpretations of the results can benefit the study.

In Finnish HLM-model the explanatory variable “Experimental approach in physics/chemistry” turned out to be a powerful predictor of science achievement in the model. Students whose physics and chemistry classes frequently involved experiments and demonstrations (18%) received 27 points higher scores than the 17% of students in whose lessons demonstrations and experiments took place only rarely.

Professor Ahtee, the Finnish science education expert commented on this result by saying that increasing the experimental approach has been a central goal for the last twenty years in science education to increase this approach. However, according to her, there is still a lot to do, especially regarding the quality aspects of experimental work.”

Another interesting Finnish finding was about “Student’s own educational expectations” The students’ own educational ambition was positively connected with their science scores in this model. The majority of Finnish students (68%) who had not made any plans beyond finishing secondary school were estimated to score 15 points lower than the 14% of students with the highest expectations, i.e. university-level education. It can be said that Finnish seventh grade students are not very career oriented, as 24% of the students remained undecided regarding their future education. The future plans of most students (68%) of this age group do not extend further than the next level of the school system, and very few students have made any further study or occupational plans at this stage. This can be considered a national characteristic, as in Finland very young students are not considered mature enough to make any final, far-reaching, trend-setting or even limiting choices regarding their future career. According to professor Ahtee, this is not necessarily a good thing. “Students are not put under any pressure, but neither do they have any goals set for them. They just go along with the flow.”

“ Student’s own educational ambition” turned out to be significant predictor of science achievement also in Hungarian model. The HLM model predicted that students who are aiming to complete university will score 46 points more than students with the lowest educational expectations or those who did not yet have any educational goals. Compared to Finnish students, Hungarians had very high educational expectations: with 58% setting their goal as finishing university and only 5% remaining undecided or choosing the lowest educational alternative (secondary school). Mr. Szalay, the Hungarian science education expert commented the results by saying that the societal changes that have occurred in Hungary since 1990 have led to increased appreciation of university degrees.

In Reinikainen’s study the results of English HLM model predicted that students who were not

living with their fathers (almost 25%) were expected to receive significantly lower scores in science than students who live with their father. The English science education expert Mrs. Clesham was not surprised by this results. According to her the rise of single parent family units and the resulting subsequent social and economic problems are significant societal issues in England at this time. There was also another interesting family related finding in English model. The English model predicts, that students with 3 or less persons living at home will score more in science than for students with households of 5 or more people. Clesham offered a number of comments on this issue. "Housing issues and the lack of physical space is a problem in England, being a relatively small country with a large population. Family units nowadays increasingly incorporate extended family i.e. grandparents. More second marriage families exist in which the children of both the first and second marriage live together. Areas with a high concentration of immigrants and refugees tend to be overcrowded in terms of living space. High property prices, particularly in the South of England explain much of the living density problem. According to Clesham this explanatory variable may also reflect poor socioeconomic status in England.

In Japan, where the living density is even higher than in England, this explanatory variable had no significance. But interesting in Japanese model was, that daily free time that student spent playing or being with friends was negatively connected with science achievement. 16% of the Japanese students did not spend any time with friends, while the most social students (25%) reported spending 3 hours or more with friends per day. For Professor Isozaki, Japanese expert of science education, the reason for this result was very clear. "In general, schools finish at 3 o'clock. Many secondary school students, especially lower secondary school students, therefore attend school clubs each day after classes until the school closes at 5 o'clock. Outside-school time is largely spent attending juku school, privately employed home-teacher lessons or doing sport. Time spent with friends is therefore taken from their studies. Where as in Japan school seemed to be very seriously taken task for students, English HLM-model was the only one where students' fun being in school was positively and significantly connected with science achievement.

Latvian and Russian models and explanations of country specific results revealed a very interesting relationship with books in these cultures. Like in every other studied country, also in Latvia, "The number of books at home" predicts student's science achievement. But what was exceptional in Latvia, was the amount of books. There were more books in Latvian homes than in any other country that participated in TIMSS-1999. Almost half of Latvian students reported to have more than 200 books at home (Martin et al., 2000a). Prof. Geske, Latvian expert of science education, explained that number of books at home variable worked even better in 1995 than today due to the changed situation in Latvia. "In soviet times books were cheap, but they were also very difficult to obtain. In Latvia, the number of books one has is

also a kind of indicator of family structure: if you live with your grandparents or parents in an old house, you are likely to have many books which the family has collected over the years.

In Russian model, “The frequency of introducing new mathematics, chemistry and physics topics from textbooks” was negatively connected to student science achievement. Dr. Kovalyova, Russian expert of science education, considered possible explanations as follows: “Russian textbooks were highly academic, not very good and not very reader friendly in times of TIMSS-1999. For example, science books had a lot of information and a lot of text, typically in very small font.” There was another interesting finding in Russian model that concerned books as well. It was about the memorization of the books. The students who disagreed most that memorization of textbooks and notes leads to high achievement in mathematics and science, were predicted to attain higher scores than those who most strongly agreed that memorization is a means of doing well in science and mathematics. A larger proportion of Russian boys (about 70%) than girls (about 55%) seemed to rely on these memorization strategies.

According to Kovalyova, “In Russia about 50% of all typical tests are (designed) just for recorded knowledge or recalling typical ways of solving typical problems. Items are neither about reasoning, nor about practical applications. On the contrary, items are just about remembering things exactly in the way they were introduced in lessons. This is a very strong tendency in Russia. It is alarming, as many textbooks present scientific facts without any questioning, practical aspects or comparisons – in an academic manner.”

The previous part of the study highlighted some of the explanatory variables of country-specific models from Reinikainen’s study with interesting emic interpretations of national science education experts. However, the author made also a summary about the existing statistically significant explanatory variables and items in the models of studied countries. He made two general observations: Firstly, it appeared that explanatory items were connected in the same direction, either positively or negatively, with science achievement in every participating country. Only the strength of connection varied, not the direction. And secondly, items that proved to be significant predictors of science achievement in a given country can actually hold any position from first to last in possible rankings of these studied countries according to the existence of its variables. This type of country ranking according to single variables or indicators is shown, for example, in the first published results of LINCAS (e.g. Martin et al., 2000a; Mullis et al., 2000b; OECD, 2001, 2004). Country rankings according to bivariate correlations between questionnaire items or indicators and student achievement might be misleading and do not provide sufficient bases upon which educational policy makers or educational practitioners can develop science education. Thus there seems to be a need for country-level models – such as HLM – which take into consideration

multivariate connections and sampling issues.

Discussion

The approach Reinikainen used in HLM-modeling on TIMSS–1999 data (cognitive results, student and school questionnaires) seemed to be most effective at describing factors connected with science achievement in England, (52% of the total variance was explained) relatively effective for Finland, Japan and Hungary (39-42%) and only moderately effective for Latvia (30%) and, especially, Russia (23%).

Reinikainen compared also the explanatory variables which proved to be statistically significant predictors of science achievement in the participating countries. The methodology applied in this study revealed that there were a total of 137 items that significantly predicted science achievement in at least one model and were also administered in each of the six countries. However, only two of these items, “Number of books at student’s home” and “Student usually does well in mathematics,” appeared to be significant predictors of science achievement in the models of the six countries. Based on the above observation, one can argue that it is extremely difficult to create a single trustworthy HLM model which could be used to predict science achievement across all participating countries, for example, across all of the 40 countries that participated in the TIMSS–1999 study.

Since the further studies of the models revealed that explanatory power of the models as well as explanatory variables differed substantially from country to country, it seems advisable to keep the international core of background questionnaires relatively small and add the number of national options instead.

Conclusion

In response to the success of the initial results of the TIMSS–1999 study in describing similarities and differences across countries, Reinikainen’s study aimed to better explain and understand the national and cultural characteristics behind science achievement in six countries. The explanatory data analysis procedures and emic interpretation of the results conducted in the study highlighted a number of issues which parents, educational practitioners, science education researchers and educational policymakers in the studied countries could focus on in their attempts to improve student achievement in science.

The use of national science education experts proved out to be very beneficial. They did not only interpret the results. They also reviewed the models and also checked the accuracy of their countries’ descriptions thus increasing the validity, reliability and accuracy of the results

of this study and also helped to avoid the “ecological fallacy” error (Boss, 2002; Crossley & Broadfoot, 1992; Postlethwaite, 1999) with the now achieved HLM models.

However, how easily these findings can be adopted in practice is another question in itself, as behind much high-profile political posturing lie deep-rooted cultural traditions and institutions that are much less amenable to change. In this sense, the author relied on science education experts in the studied countries in the hope that they may be active missionaries of the results of this study. Changes in educational traditions cannot be enforced by outsiders to the culture. These changes are more likely to become a reality if people of influence in the field in the respective countries are actively involved in the process. Reinikainen’s study showed that cross-national research into school achievement can reveal more than a simple ranking of nations according to science achievement or strengths of certain variables or indicators. Comparative studies of achievement provide us with a lens through which we can view culture in action and, most importantly, they provide educational practitioners a good basis on which to co-operate for the benefit of our children.

References

- Alexander, R. (1999). Culture in pedagogy, pedagogy across cultures. In R. Alexander, P. Broadfoot & D. Philipps (Eds.), *Learning from comparing: New directions in comparative educational research* (Vol. 1, 149–180). Oxford: Symposium Books.
- Bempechat, J., Jimenez, N., & Boulay, B. (2002). Cultural-cognitive issues in academic achievement: New directions for cross-national research. In A. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 117–149). Washington, DC: National Academy Press.
- Bos, K. Tj. (2002). *Benefits and limitations of large-scale international comparative achievement studies. The case of IEA’s TIMSS*. Thesis University of Twente, Enschede: PrintPartners Ipskamp.
- Creswell, J. W. & Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, London, New York: Sage Publications.
- Crossley, M., & Broadfoot, P. (1992). Comparative and international research in education: scope, problems and potential. *British Educational Research Journal*, 18(2), 99–112.
- Cogan, L. S., & Schmidt, W. H. (1999). Middle school maths reform. *Middle Matters*, 8 (Fall), 2–3.
- Gonzalez, E. J., & Miles, J. (Eds.). (2001a). *TIMSS 1999 user guide for the international database*. MA: International Study Center, Lynch School of Education, Boston College.
- Gonzalez, E. J., & Miles, J. (Eds.). (2001b). *TIMSS 1999 user guide for the international database. Supplement 1. International version of the background questionnaires*. MA: International Study Center, Lynch School of Education, Boston College.
- Gonzalez, E. J., & Miles, J. (Eds.). (2001c). *TIMSS 1999. User guide for the international database*.

- Supplement 2. Documentation of national adaptation of background questionnaire items.* MA: International Study Center, Lynch School of Education, Boston College.
- Kaiser, G. (1999). International comparisons in mathematics education under the perspective of comparative education. In G. Kaiser, E. Luna & I. Huntley (Eds.), *Studies in Mathematics Education* (pp. 3–15). London: Falmer Press.
- Kaiser, G., Leung, F. K. S., Romberg, T., & Yaschenko, I. (2002). International Comparisons in Mathematics Education: An Overview. *ICM 2002*, 1, 631–646.
- Keitel, C. & Kilpatrick, J. (1999). The rationality and irrationality of international comparative studies. In G. Kaiser, E. Luna and I. Huntley (Eds.), *International Comparisons in Mathematics Education*. (pp. 241–256). Falmer Press.
- 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.
- Kyriakides, L., & Charalambous, C. (2004). Extending the scope of analyzing data of IEA studies: Applying multilevel modeling techniques to analyse TIMSS data. In C. Papanastasiou (Ed.), *Proceedings of the IRC–2004. TIMSS* (Vol. 1, 69–86). Cyprus University, Nicosia.
- Martin, M. O., Mullis, I. V. S., Gonzalez, Gregory, K. D., Smith, T. A., Chrostowski, S. J Garden, A. G., & Kelly, D. L. (2000a). *TIMSS 1999 International Science Report. Findings from IEA’s repeat of the Third International Mathematics and Science Study at eighth grade.* MA: Boston College, Lynch School of Education.
- Martin, M. O., Mullis, I. V. S., Gregory, K., Hoyle, C., & Shen, C. (2000b.) *Effective schools in science and mathematics. IEA’s Third International Mathematics and Science Study: TIMSS.* MA: International Study Center, Boston College, Chestnut Hill, USA.
- Moghaddam, F. M., Walker, B. R., & Harré, R. (2003). Cultural distance, levels of abstraction, and the advantages of mixed methods. In A. Tashakkori & C. Teddlie. *Handbook of mixed methods in social & behavioral research* (pp. 111–134). Thousand Oaks, London, New Delhi: Sage Publications.
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie. *Handbook of mixed methods in social & behavioral research* (pp. 189–208). Thousand Oaks, London, New Delhi: Sage Publications.
- Mullis, I. V. S., Martin, M. O., Stemler, S. E, Gonzalez, E. J., Grogory, K. D., O’Connor, K. M., Crotowski, S. J., & Smith, T. A. (2000). *TIMSS 1999 international mathematics report. Findings from IEA’s repeat of the Third International Mathematics and Science Study at the eight grade.* MA: Lynch School of Education, Boston College.
- OECD (2001). *Knowledge and skills for life. First results from the OECD Programme for International Student Assessment. PISA 2000.* Paris: OECD.
- OECD (2004). *Learning for Tomorrow’s World. First results from the PISA 2003.* Paris: OECD
- Pepin, B. (2003). *Can we compare like with like in comparative educational research? On the*

- importance of studying anomalies*. Retrieved September, 2006, from:
http://scholar.google.com/scholar?hl=en&lr=&q=cache:_1VMpMRGX7gJ:www.lehrplan.ch/Network/pdf/PepinACER.pdf+related:_1VMpMRGX7gJ:scholar.google.com/.
- Postlethwaite, T. N. (1999). *International studies of educational achievement: methodological issues*. Hong Kong: Comparative Education Research Centre, University of Hong Kong.
- Reinikainen, P. (2007). *Sequential Explanatory Study of Factors Connected with Science Achievement in Six Countries: Finland, England, Hungary, Japan, Latvia and Russia. Study based on TIMSS 1999*. Doctoral thesis. Institute for Educational Research, University of Jyväskylä, Jyväskylä.
- Schmidt, W. H., & Burstein, L. (1993). Concomitants of growth. In L. Burstein (Ed.), *The IEA Study of Mathematics III: Student growth and classroom Processes* (pp. 309–327). Oxford: Pergamon Press.
- Schmidt, W. H., & McKnight, C. C. (1998). What can we really learn from TIMSS? *Science*, 282(5395), 1830–1831.
- Sjøberg, S. (2003). Science and technology education current challenges and possible solutions. In E. Jenkins (Ed.), *Innovations in science and technology education*. (Vol VIII). Paris: UNESCO.
- Stevenson, H. W. (1998). A TIMSS primer. Lessons and implications for U.S. Education. *Fordham Report*, 2(7).
- Tashakkori, A. & Teddlie, C. (Eds.). (2003). *Handbook on mixed methods in the behavioral and social sciences*. Thousand Oaks, CA: Sage Publications.

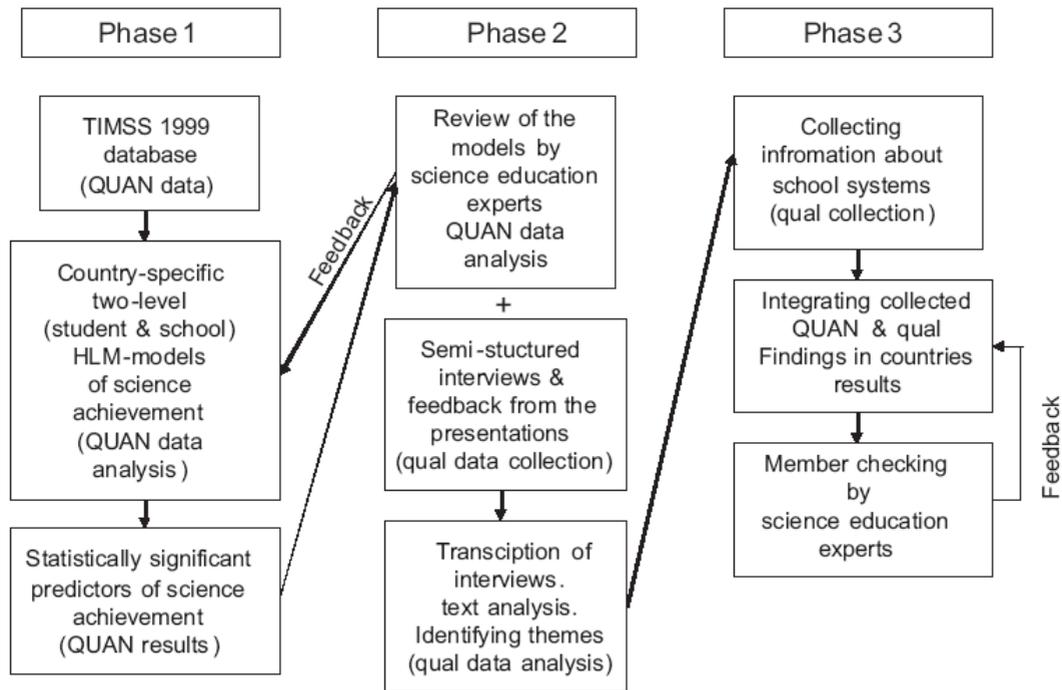


Figure 1. Visual diagram of the overall procedures in this study using qualitative data to explain quantitative results. The notation system (Morse, 2003) uses pluses (+) to indicate methods that occur at the same time and arrows (→) to indicate methods that occur in sequence. The use of parentheses to indicate the notation system designates the relative importance of the methods. The primary method is indicated using uppercase letters and the secondary with lowercase letters.

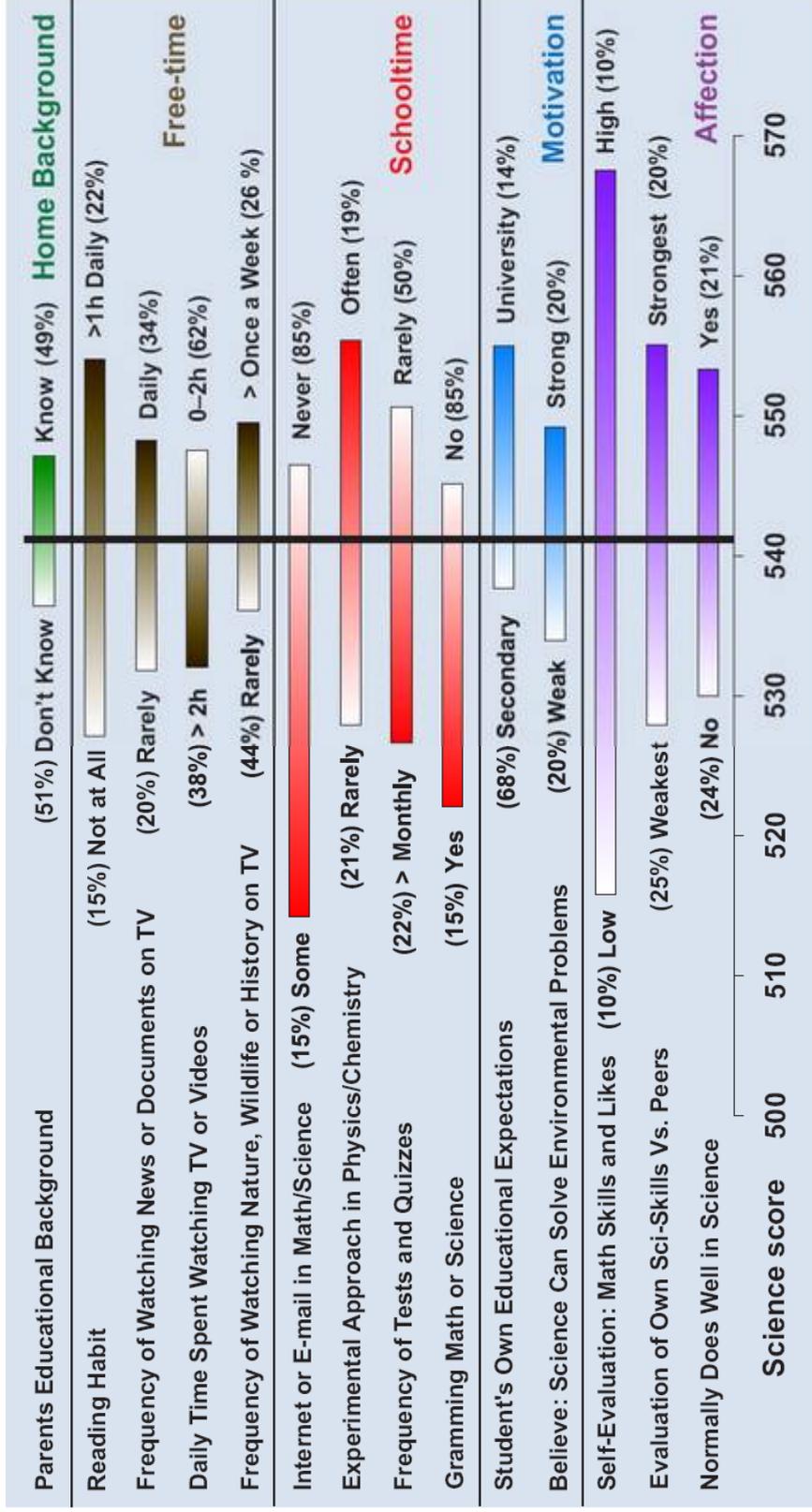


Figure 2 Factors connected with science achievement in Finland: explanatory variables, their explanatory power and distribution.

Graph description: the black text introduces each explanatory variable. The names of the variables are abbreviated and simplified. The vertical line represents the average science score (541 points) of the Finnish students included in the model. The length of the horizontal bars indicates the explanatory power of the respective variables. The longer the bar, the more it explains the variation. The different bar colors indicate the grouping of the variables. The bar shift gradually from lighter to deeper shades. A lighter shade at the end of the bar represents lesser values of variables or factors (e.g., infrequent occasions, negative attitudes, etc.) and a darker shade correspondingly indicates higher values of variables or factors (e.g., frequent occasions, positive attitudes). The ends of the bars are labeled with the values of the respective variable. The original scales used in the background questionnaires (see Gonzalez et al., 2001a, 2001b, 2001c) were rescaled, condensed and simplified when used in the model. The percentage of students at the end points of the used scale is marked in parentheses. If the percentages of the bar total 100%, the scaling of the variable has been dichotomized. This is the case, for example, with the first explanatory variable Parents educational background. Of the Finnish eighth-graders, 51% did not know the educational background of their parents. The model gives them an estimation of scoring around 11 points lower in science than the remaining 49% of the students, who were able to provide this information. The second bar can be interpreted as “the least frequent readers of books”, i.e. the 15% of students who read no books at all are estimated to score 14 points lower in science than average and as much as 27 points lower than the most frequent readers, i.e. the 22% of students who read at least one hour per day.