VILNIUS UNIVERSITY
FACULTY OF MATHEMATICS AND INFORMATICS
MODELLING AND DATA ANALYSIS
MASTER'S STUDY PROGRAMME

Self-Regulated Learning and it's Effect on Reading Achievement: Analysis of Lithuania's Participation in PISA 2018

Master's thesis

Author: Karina Silkina
VU email address: karina.silkina@mif.stud.vu.lt
Supervisor: prof. Rimas Norvaiša

Vilnius

2023

Savarankiškas mokymasis ir jo įtaka skaitymo pasiekimams: Lietuvos dalyvavimo PISA 2018 analizė

Santrauka

Raktiniai žodžiai: (skaitymo raštingumas, savivaldus mokymasis, hierarchinis tiesinis modelis, duomenų analizė)

Šio darbo tikslas buvo ištirti PISA 2018 metų Lietuvos duomenis, pritaikyti savivaldaus mokymosi arba savireguliacijos teoriją ir įvertinti savireguliacijos įgūdžių įtaką ir reikšmingumą skaistymo raštingumui Lietuvoje. Tikslui pasiekti buvo pasitelktas dviejų lygių hierarchinis tiesinis modelis. Modelio rezultatai buvo palyginti su analogiškais modeliais Latvijai ir Estijai. Rezultatai parodė tam tikrų savireguliacijos įgūdžių reikšmingumą skaitymo raštingumui ir pabrėžė Lietuvos skirtumus nuo Latvijos bei Estijos. Analizės rezultatai gali paskatinti toliau tirti aspektus, lemiančius savireguliacijos įgūdžių skirtumus bei kodėl Lietuvos moksleiviai šioje srityje yra silpnesni.

Self-Regulated Learning and it's Effect on Reading Achievement: Analysis of Lithuania's Participation in PISA 2018

Abstract

The aim of this thesis was to investigate the Programme of International Student Assessment (PISA) 2018 data from Lithuania, apply self-regulated learning theory to the data, and evaluate the effect and significance of self-regulated learning skills on reading literacy in Lithuania. This involved building a two-level hierarchical linear model and comparing the results obtained in Lithuania with those from Latvia and Estonia. Results revealed significance of certain self-regulated learning elements for reading literacy, and highlighted Lithuania's differences from Latvia and Estonia. The findings of this analysis may lead to further investigation into the specific factors that contribute to the differences between Baltic countries in terms of self-regulation, particularly why Lithuanian students tend to be weaker in this area.

Keywords: (reading literacy, self-regulated learning, hierarchical linear modelling, PISA, data analysis)

Contents

1 Introduction

The COVID-19 pandemic has presented numerous challenges to children's education, including the sudden shift to remote learning. This sudden transition has highlighted the vulnerability and lack of preparedness of educational systems in the face of unexpected changes, leading to negative impacts on the educational environment, content, and nature. As a result, there is a growing interest in research to understand the effects of these changes on societies, educational systems, and individuals, in order to better prepare for potential future crises.

In Lithuanian context, an extensive project was carried out to assess and analyse the consequences of compulsory remote learning imposed during quarantine in Lithuania on children's education and health [33]. Results of the project indicates decreased motivation to learn, lack of self-learning skills, lack of digital competence (for students and parents as well), lack of appropriate study conditions, etc [33]. Participants of the project concludes that remote learning revealed both weaknesses of as well as possibilities for the education system. Some suggestions for policy makers were also proposed however, even though the potential of developing self-learning skills was acknowledged [33], this was not presented as the topic worth paying attention to in the future. Unconventional settings, such as remote learning during the COVID-19 pandemic, when students had to quickly organize and self-regulate their learning with little preparation [10], highlight the importance of fostering a more proactive student role in the learning process. While most would agree that in-person learning is a priority, the current post-pandemic context and its outcomes suggest that the development of students' skills to study independently, as well as an examination of how these skills affect student achievement, should be considered important topics for future research.

Self-regulated learning (SRL) has long been recognised as an important contributor to learning success in various educational contexts, whether online or offline [10, 7, 8]. Even though nowadays there is a large variety of different theoretical constructs and definitions of this concept, most definitions agree on the common ground and view self-regulated learning students as metacognitively, motivationally and behaviourally active participants in their learning process who self-regulate their actions to achieve specific goals [5, 21, 7]. It is a very broad conceptual framework which consists of different strategies and components. Additionally, there are numerous studies which support the connection between these strategies and academic achievement [8]. For example, there is evidence that different learning strategies differently affect reading comprehension and achievement [14, 36, 11, 15], as well as learning outcomes in other areas, such as chemistry [34]. Other studies demonstrate the importance of motivational and behavioural SRL aspects [10, 30, 12, 28, 19].

Programme of International Student Assessment (PISA) is an ongoing worldwide programme which measures 15 year old student's performance in three domains: mathematics, science, and reading. The main aim of this assessment is to measure an extent to which students have acquired the knowledge and skills of each domain that are essential for full participation in modern societies [23]. Assessment is carried out in cycles every three years and in each round, one of the three topics is tested in more detail. In 2018 main domain of the assessment was reading literacy. According to assessment's analytical framework, individual's reading practices, motivation, attitudes, and awareness of effective reading strategies all contribute significantly to an individual's reading ability. Students who read frequently, are interested in reading, feel confident in their abilities, and know how to use strategies such as summarizing or searching for information tend to be more proficient in reading. These practices, motivation, and metacognition are not only potential predictors of reading achievement, but also important goals or outcomes of education that can drive lifelong learning [23].

Hence, the main goals of this thesis are as follows:

- Investigate dataset of PISA 2018 and find out which variables are measuring different self-regulated learning components

- Apply self-regulated learning theory on PISA 2018 data and evaluate the effect and significance SRL skills had on reading achievement in Programme of International Student Assessment 2018 in Lithuania by building a two-level hierarchical linear model.

- Compare Lithuanian obtained results with models from Latvia and Estonia.

Final results of this work revealed that use of specific learning strategies is significant predictor of reading literacy in all three countries. Feelings of competence was a significant factor in reading literacy in Lithuania as well as Latvia and Estonia, with self-efficacy having no significant effect. Enjoyment of reading was found to be a significant predictor in all three countries, with the effect being stronger in Latvia and Estonia. Negative emotions had a complex relationship with reading literacy, with the perception of difficulty having a negative impact and fear of failure having a positive impact, particularly stronger in Latvia and Estonia. The economic, social, and cultural status of a school had a stronger impact on student performance in Lithuania. Overall, Lithuanian students appeared to be weaker in terms of self-regulated learning compared to Latvian and Estonian students and this gives more possibilities for future research.

Next part of the thesis outlines the theoretical framework while the following parts describe data and methodology used. Finally, the thesis will be concluded by exploratory analysis and model description together with explained results, conclusions and future implications.

2 Theoretical framework

The theory of self-regulated learning (SRL) developed as a result of the shift towards the belief that students should actively participate in their own learning process, rather than simply following instructional theories [7]. SRL as a concept has a number of different theoretical definitions and models but all of them share common assumptions and features [29]. From a social cognitive perspective, self-regulated learning is seen as the interaction between personal, behavioural, and environmental processes [1, 4]. It refers to self-generated thoughts, feelings and actions that are planned and adapted for the attainment of personal goals [4]. In self-regulated learning, students are expected to take an active role in the learning process and engage in it willingly, rather than simply responding to teaching [9, 3, 4]. In other words, self-regulated learning involves developing control over the educational process and taking responsibility for one's own learning in order to achieve specific goals. This involves the student setting learning goals and taking charge of the learning process in order to achieve them [9].

Paul R. Pintrich, an educational psychologist and researcher, developed a more structured theoretical framework for self-regulated learning that combined various concepts associated with SRL into a single framework [29]. According to Pintrich, self-regulated learning *"is an active, constructive process* *where learners set their goals for learning and afterwards attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by goals and the contextual features in the environment"* [29]. Activities coming from the definition can mediate relationships between individuals, contextual environment and overall achievement [29]. According to Pintrich's framework [29], there are four main processes involved in self-regulated learning: goal-setting, monitoring, control, and reflection. These processes take place across different areas of regulation, such as cognition, motivation, behaviour, and context. The phases of this framework do not have to occur in a specific order or be hierarchical in structure. In many models and real-life examples, these stages can occur simultaneously [29]. Therefore, this general framework offers a taxonomy for the processes and components of self-regulated learning, which can be used to organize research in a structured manner.

This work adapts and applies two main areas and components of the above described framework - regulation of cognition and motivation. Both these areas are in detail described in a further subchapters.

Self-regulation of cognition

Cognitive strategies are tools that students can use to learn and acquire knowledge more efficiently and they are a key part of self-regulated learning [30]. Some examples of strategies for learning and retaining information include memorization, note taking, summarizing, and similar techniques. These strategies can range from simple memory aids to more complex methods for organizing and synthesizing information [29]. According to Pintrich's framework, different learning strategies can be used as part of a cognitive control process in which individuals evaluate their thought process and decide whether to continue using a particular strategy or switch to a different one. These activities involve actively regulating and modifying one's own cognition in order to better understand and retain information [29]. Therefore, process of selecting and implementing strategies for memory, learning, and problem solving is a key aspect of the cognitive control process. This involves making decisions about which strategies to use and how to use them effectively in order to achieve one's learning and problem-solving goals [29].

Empirical evidence has shown that selection of appropriate cognitive strategies can have a positive influence on learning results and performance. For example, an analysis of PISA 2009 data from Shanghai found that metacognitive strategies, such as understanding and remembering, as well as summarizing, were significantly related to students' reading literacy [36]. These findings were also supported by an analysis of 2018 data conducted in Greece [14]. Similarly, using data from Turkey in 2009, researchers found that higher-level control and elaboration strategies (as opposed to simple memorization) had a direct impact on students' enjoyment of reading and their reading scores [11].

Studies mentioned above and many other have found that students who use strategies such as understanding and summarizing texts tend to have better reading comprehension and academic achievement. While the effectiveness of these strategies may vary from country to country, the overall trend is clear: using these strategies can have a positive impact on learning outcomes.

Elements of motivation

Simply being aware of various learning strategies is often not sufficient to promote self-regulation and improve academic achievement. If an individual lacks the motivation to use their self-regulation skills, these skills are of little value [4, 30, 29]. Hence, it is also believed that students' motivational attitudes and beliefs are related to their academic performance and achievement [30]. Pintrich's framework for self-regulated learning incorporates a general model of motivation that includes expectancy, value, and affective components. This model is known as the expectancy-value model of motivation [30].

Expectancy component of student motivation includes students' belief that they are capable of performing a task [30]. Self-efficacy judgments, or beliefs about one's own ability to perform a task, can be modified at any point during the learning process based on actual performance, feedback, and efforts to regulate these beliefs about one's competence [30]. People who are self-doubters are more likely to withdraw if they do not believe in their competence to perform a certain task [4]. The more capable people believe themselves to be, the higher goals they usually set for themselves hence the more committed they remain while trying to achieve them [4]. There is also empirical evidence to support these assumptions. For example, research conducted in among Austrian secondary school students showed that students who believe they are competent tend to be better at managing their time, setting goals, using metacognitive strategies, and have higher levels of intrinsic motivation [10]. Results from questionnaire administered to seventh graders in United States revealed that self-efficacy is positively related not only to cognitive engagement but to academic performance as well [30].

The value component of student motivation is about the student's goals and interest in the task [30]. It is believed that students who are more interested in performing a task are more likely to engage in the metacognitive activity and strategy use which in turn improves their results [30]. Interest in a specific task is linked to increased learning, persistence and effort [29]. Some studies suggest that motivation is one of the most important components of SRL explaining student performance. For example, Hong Kong study on PISA 2009 concluded that motivation is the most crucial SRL factor which explains strong performance of Hong Kong students [20]. In a similar way, Chinese study on 2018 data demonstrated that achievement goals contribute to academic performance as well [31]. Finally, similar findings were presented in already mentioned Greek study [14]. Additionally, other evidence shows that intrinsic value of a task has a positive effect on cognitive engagement as well as performance [30, 14], students who use different learning strategies for reading related tasks are more likely to enjoy reading as a whole [36] which in turn improves their reading performance [11].

Affective components of motivation describe emotional reaction to a task [30]. One of the examples in school learning contexts could be test anxiety, which was found to be a strong predictor of performance among US seventh graders [30, 29]. However, the student's reactions and its relationship with achievement is not straightforward. For example, it was observed that students, who are more anxious before tests, can be as persistent as students with lower anxiety, however they often did not use appropriate learning strategies [30]. Another affective could be perception of how difficult a particular task is and how it affects task performance [29]. Generally, negative emotions about a specific task are considered to negatively impact performance

To summarize assumptions coming from theoretical framework described above, following hypotheses were raised:

H1: Students who employ various metacognitive strategies and find them helpful are more likely to achieve higher score in reading literacy.

H2: Students who believe in they competence and are self-efficacious are more likely to achieve higher score in reading literacy.

H3: Students who are motivated to master tasks, are mastery goal oriented and enjoy reading are more likely to achieve higher score in reading literacy.

H4: Students who have negative reactions about specific tasks are more likely to be less successful in reading literacy.

The whole framework described and which will be analysed in this work is also summarised by Figure 1.

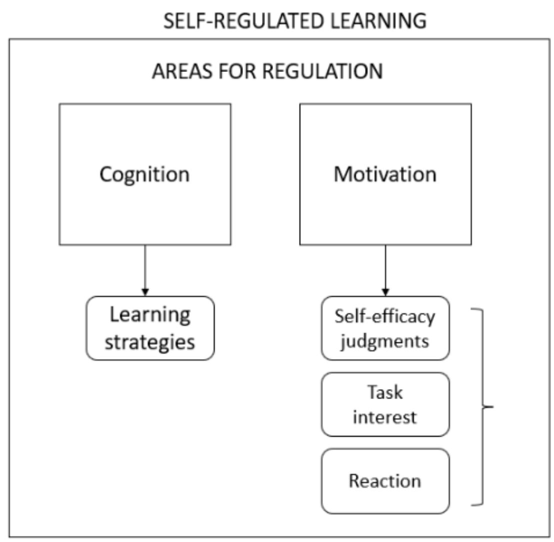

Figure 1: Self-regulated learning framework

3 Data description

The OECD Programme for International Student Assessment (PISA) is an ongoing global assessment conducted every 3 years that aims to assess knowledge and skills of 15-year-old students and how prepared they are to participate in contemporary societies [23]. PISA evaluates not only student's ability to recall specific information, but also their ability to use what they have learned to solve novel problems and adapt to new situations [23].

Assessment focuses on three main subjects – reading, mathematics and science – and in each cycle one of the subjects is tested in more detail. In 2018 the survey focused on reading, while mathematics and science were assessed as minor domains [23]. The reading assessment included a computer-based test that featured a combination of multiple-choice questions and open-response questions requiring students to generate their own responses. Included questions measured students' ability to understand, use, reflect on, and engage with written texts in various formats, such as literary and informational texts, and to think critically and creatively. In addition to the assessment test, students and school principals also completed a series of questionnaires. The student questionnaire gathered information about the student's personal characteristics, home life, school and learning experiences. Students were also asked about their reading habits and attitudes, as well as about the reading-related support and resources available to them at home and at school. The school questionnaire collected data about the school system and learning environment [23].

Further subsections describe main characteristics of 2018 dataset and all of the variables which will be included in the subsequent analysis.

3.1 Data characteristics

The PISA methodology includes several key components that are essential for accurate analysis of the survey results. One important aspect is the sampling design, which refers to selecting a representative group of individuals from the population being studied. Sample weights are also an important part of the methodology, as they are used to adjust the results of the survey to account for any biases or imbalances in the sample. Plausible values are another key component of the methodology, which are derived from the data collected in the survey and are used to improve the accuracy of the estimates produced. All of these components work together to ensure that the survey results are reliable and accurately reflect the broader population being studied [22]. All three components mentioned are described in more detail below.

Sampling design. Instead of simple random sampling, students for the survey are sampled in two stages [22]. The first-stage samples individual schools that had or potentially could have 15-year-old students at the time of assessment [22, 26]. Before selecting the schools for the sample, they are divided into distinct groups called explicit strata based on specific characteristics of the school. These were created in order to increase the accuracy of estimates based on a sample [26]. Strata may vary from country to country, but they usually include different regions and types of schools [16, 26]. Schools are chosen randomly from within each of these specific strata, with the probability of selection being proportional to the size of the school [16]. Within the selected schools, a simple random sample of students is then drawn. In PISA, typically 30 or 35 students from the population of 15-year-olds are randomly selected from within the selected schools [16, 22]. This type of survey design has important implications for how the data is analysed by secondary users [16].

Sampling weights. In order to accurately represent the full population of students in the PISA study, the final sample of students from each country/economy is chosen randomly, but the selection probabilities of individual students may vary. To correct for this, survey weights are applied to ensure that their contribution to the overall population estimate is proportional. These weights, called sampling weights, are used to control the influence of each student on the final analysis [26, 16]. If the appropriate weights are not applied to the data, the characteristics of some students or schools may be overrepresented or under-represented in the analysis, potentially resulting in biased estimates. It is important to apply weights to ensure that the analysis accurately reflects the full population being studied [16].

Plausible values. Rather than providing a single measure to report on student achievement, PISA survey includes several plausible values [17, 22]. In PISA 2018 10 plausible values for reading literacy and other domains were reported. In other words, 10 separate variables are included in the final dataset to measure the achievement of the particular domain. Each plausible value represents a randomly drawn estimate from a distribution of possible values (posterior distribution) for that student and they are used to account for uncertainty in score estimation [17]. When conducting an analysis using plausible values, the statistical model or analysis is conducted separately for each value and then results are combined. It is not recommended to take the average of the plausible values for use in analysis [17, 22].

More details on how these components are handled practically during the subsequent analysis will be provided in a further section dedicated to methodology.

3.2 Key variables

This subchapter details each of the key dependant and independent variable included in the analysis. The selected variables are combination of different measures collected from reading tests and contextual questionnaires.

Reading literacy as dependent variable

Since reading is a major domain in 2018's assessment, reading literacy (or reading proficiency) was selected as a dependent or outcome variable in this analysis. According to 2018 assessment's analytical framework, reading literacy is *"understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society"* [23]. The reading literacy assessment administered by PISA evaluates the ability to read and comprehend various types of texts and complete tasks of varying difficulty levels [24]. Final scores for reading literacy are provided as a set of 10 separate plausible values (as *PV1READ, PV2READ* and etc.).

Independent variables

There are two main types of independent variables used in this analysis: student level and school level. Student level variables refer to characteristics related to self-regulated learning, specifically learning strategies and motivational elements. School-level variables, on the other hand, refer to characteristics of the school or educational setting in which survey was taking place.

Learning strategies. 2018 survey included to scenarios which were assessing students cognition and their use of learning strategies. Two types of strategies were assessed: understanding and remembering, and summarising. Students were asked to evaluate the effectiveness of different strategies for completing a reading task [26]. These strategies were also rated by reading experts through a series of pairwise comparisons. The resulting hierarchy of strategies for each task was determined based on the consensus of at least 80% of the experts [26]. Using the hierarchy of strategies generated by the expert ratings, rules were established to calculate a score for each student based on how often they chose a more effective strategy over a less effective one [26]. The final derived scores ranges from 0 to 1. The higher the score, the higher the number of times in which a student chose an expert-validated strategy over a less useful one [26]. Therefore, the following variables were obtained:

- Understanding and remembering (*UNDREM*)
- Summarising (*METASUM*)

Motivational elements. Variables attributed to motivational elements are part of scale indices, which were constructed through the scaling of multiple items (questions) [25]. The following motivational elements were selected:

- Perception of competence in reading (*SCREADCOMP*) - included three questions about student's self-perception as a reader. Positive values of the index indicates greater perception of competence than the OECD average [25].
- Self-efficacy (*RESILIENCE*) - index concerns student's self-efficacy and belief in their own capabilities. Positive values of this index mean that student reported higher self-efficacy than did average student across all countries participating in the assessment [25].
- Perception of difficulty (*SCREADDIFF*) - index consists of questions whether student encountered any difficulties while performing reading tasks. Positive values represent a higher perception of difficulty compared to the average of the OECD countries [25].
- Fear of failure (*GFOFAIL*) - students were asked if they experience feelings of fear or anxiety of not being successful. Positive values indicate greater fear of failure than did the average student across OECD countries [25].
- Enjoyment of reading (*JOYREAD*) - index was created based on questions about whether reading is a hobby or enjoyable activity for the students. Values above zero indicate that student enjoys reading to greater extent than the average student across participating countries [25].
- Motivation to master tasks (*WORKMAST*) - index was composed of questions about whether students find satisfaction in completing tasks and whether they are motivated to finish them. Positive values indicate greater motivation [25]
- Learning goals (*MASTGOAL*) - index for learning goals was constructed by asking questions about student's goals and how ambitious they are. Positive values show more ambitious learning goals than the average student across OECD countries [25].

Additionally, to control for gender effect, student-level gender variable was included (*MALE*) derived from the questionnaire and converted to a binary form.

School-level variables. Three variables from a questionnaire administered to school principals were included to control for school-level effect:

- School's economic, social and cultural status (*SCHOOL_ESCS*) - average of student's individual ESCS across a school.
- A type of area in which school is located (*WHICH_AREA*) - possible values: rural area, small town, town, city or a large city.
- Public or private school (*PUBLIC*) - binary variable, where 0 means that school is public and 1 - private.

4 Methodology

To address the research questions and achieve the objective of this work, hierarchical linear modelling approach will be used. Hierarchical linear models (HLM) are used to analyse data that have a nested or hierarchical structure. This means that the data can be organised into multiple levels, with lower levels nested within higher levels. This type of data structure is common in many fields, especially in education. Considering a hierarchical model when analysing educational data is advisable due to the fact that, for example, students within the same school tend to be more similar to each other than students attending different schools. This can help to account for shared factors within a school that may affect student outcomes [18]. Hierarchical linear models allow researchers to take into account the nested structure of the data and to estimate the effects of variables at each level of the hierarchy while controlling for the other variables at that level and all higher levels [13].

In this thesis, a two-level hierarchical linear modelling approach will be used, with first level representing students and second level representing schools. Typically, HLM is performed in three main steps [35]. The first step is to determine whether hierarchical structure of the data should be considered. This can be achieved with unconditional (null) model where no student or school characteristics are included. Mathematical form for unconditional model is as follows:

$$Y_{ij} = \gamma_{00} + U_{0j} + r_{ij}, \quad i = 1, ..., n_j, \quad j = 1, ..., J. \tag{1}$$

Where Y_{i_j} represents outcome, γ_{00} represents the intercept, $Var(r_{ij}) = \sigma^2$ - within group variance and $Var(U_{0j}) = \tau_{00}$ - between group variance [13, 17]. In other words, a one-way analysis of variance is conducted to determine if the variability in the outcome variable Yij (for person i in school j) significantly differs from zero at different levels of analysis. This tests for the existence of differences in the outcome variable at the group level [13].

Unconditional model is also used to compute the interclass correlation coefficient (ICC) which has the following formula:

$$ICC = \tau_{00}/(\sigma^2 + \tau_{00}) \tag{2}$$

Here τ_{00} is equal to school-level variance while σ^2 is a student-level variance. Coefficient can be interpreted as a proportion of variation at level-2 for the given outcome measure [17]. In the context of the current analysis, it's a proportion of variability in reading literacy at the school level. If, for example, this measure would be equal to 0, this would mean there is no variation on a school level hence hierarchical modelling approach is not necessary.

After evaluating unconditional model, separate level-1 models are developed for each level-2 unit. This type of a models are also called within unit models as they describe the effects in a context of a single group [13]. On level-2, level-1 regression coefficients are used as outcome variables and are related to each of the level-2 predictors [13]. The model has the following form [13]:

Level-1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij} \tag{3}$$

Level-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}G_j + U_{0j} \tag{4}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}G_j + U_{1j} \tag{5}$$

Here:

Y_{ij} - dependent variable which is measure for the ith level-1 unit nested within jth level-2 unit;

X_{ij} - level-1 predictor;

β_{0j} - intercept for jth level-2 unit;

β_{1j} - regression coefficient associated with the predictor or slope for the jth level-2 unit;

G_j - value on the level-2 predictor;

r_{ij} - random error related to ith level-1 unit nested within jth level-2 unit;

γ_{00} and γ_{10} - overall mean intercepts adjusted for G;

γ_{01} and γ_{11} - regression coefficient associated with G relative to level-1 intercept and slope accordingly;

U_{0j} and U_{1j} - random effects of the jth level-2 unit adjusted for G on the intercept and slope accordingly.

Finally, combined two-level model is derived, which contains predictors from both levels. This model can include both fixed and random effects [13]. Generally, random effect is any variable where main interested lays in the distribution of the outcome across all categories, which are modelled as fixed effect [17]. In the case of this analysis, only the intercept is allowed to vary. Finally, combined model for this particular analysis considering key variables described in a previous chapter will have a following form:

$$\begin{aligned} Y_{ij} = \gamma_{00} + \gamma_{01}SHCOOL\_ESCS_j + \gamma_{02}WHICH\_AREA_j + \gamma_{03}PUBLIC_j + \\ \gamma_{10}MALE_{ij} + \gamma_{20}UNDREM_{ij} + \gamma_{30}METASUM_{ij} + \\ \gamma_{40}SCREADCOMP_{ij} + \gamma_{50}RESILIENCE_{ij} + \gamma_{60}SCREADDIFF_{ij} + \\ \gamma_{70}GFOFAIL_{ij} + \gamma_{80}JOYREAD_{ij} + \gamma_{90}WORKMAST_{ij} \\ \gamma_{100}MASTGOAL_{ij} + U_{1j}MALE_{ij} + U_{2j}UNDREM_{ij} + U_{3j}METASUM_{ij} + \\ U_{4j}SCREADCOMP_{ij} + U_{5j}RESILIENCE_{ij} + U_{6j}SCREADDIFF_{ij} + U_{7j}GFOFAIL_{ij} + \\ U_{8j}JOYREAD_{ij} + U_{9j}WORKMAST_{ij} + U_{10j}MASTGOAL_{ij} + U_{0j} + r_{ij}. \end{aligned} \tag{6}$$

The fact that the outcome variable is not an assigned single score but rather a set of 10 plausible values adds additional complexity to the analysis. It is important to carefully consider the correct approach for analyzing data with plausible values, as common mistakes such as using only one value or the mean of the plausible values as a single estimate of achievement can lead to underestimation of the standard errors of estimated statistics [2]. Therefore, in this study, analysis will be conducted separately for each of the 10 plausible values [2, 17]. To compute the point estimate of a specific statistic, statistic will be computed once for each plausible value and then the average will be taken. Therefore, the same model will be estimated 10 times.

Another thing to consider before conducting the hierarchical analysis, is the use of sampling weights. PISA 2018 provides both school-level w_i and student-level w_{ij} weights that can be used to correctly analyse specific populations [18]. However, when analysing data from multiple levels at the same time, these weights need to be used or adapted differently in order to accurately account for the hierarchical structure of the data [18]. Using only the final student weight is not sufficient for multilevel analysis and using unscaled student weights may result in biased variance estimates [18]. Hence, to properly incorporate weights into hierarchical analysis, it may be necessary to scale the level one weights.

In this analysis, final student weights were scaled using cluster weights approach. This approach was found to be one of the least biased across different options [18, 2]. Here, on level-1, weights are scaled up to add up to the cluster size [18, 2]. Cluster weight for student i in school j was computed in the following way:

$$W_{ij}^{*} = W_{ij}\frac{n_j}{\sum_j W_{ij}} \tag{7}$$

Here W_{ij} is the total student weight, n_j is the student sample size in school j and $\sum_j W_{ij}$ is the sum of total student weights in school j [2]. For the second school-level, available in the dataset school weights will be used.

Finally, analysis will be conducted using *WeMix* package from statistical software *R*, which is used to to fit hierarchical linear models. This package was found to be more beneficial in comparison with other due to it's ability to apply sampling weights at multiple levels of the analysis [27].

5 Results

5.1 Exploratory analysis

To begin with, general descriptive statistics are calculated from the data. In Lithuania, total sample size was 6885 observations with 362 schools. Mean reading literacy score is equal to 475.87 (see Table 1). In comparison to other Baltic countries, Latvia has a mean reading literacy score of 478.7 based on a sample of 5303 individuals from 308 schools. Estonia, on the other hand, had the highest mean reading performance score of 523.02, with a sample size of 5316 individuals from 230 schools (see Table 1).

Country	Freq	Mean	SE	SD	SE
LTU	6885	475.87	1.52	94.3	1
LVA	5303	478.7	1.62	90.03	1.07
EST	5316	523.02	1.84	93.21	1.18

Table 1: Mean achievement score and standard error in each country

Graphical representation of mean reading literacy score differences in each school and each country can be observed in Figure 2.

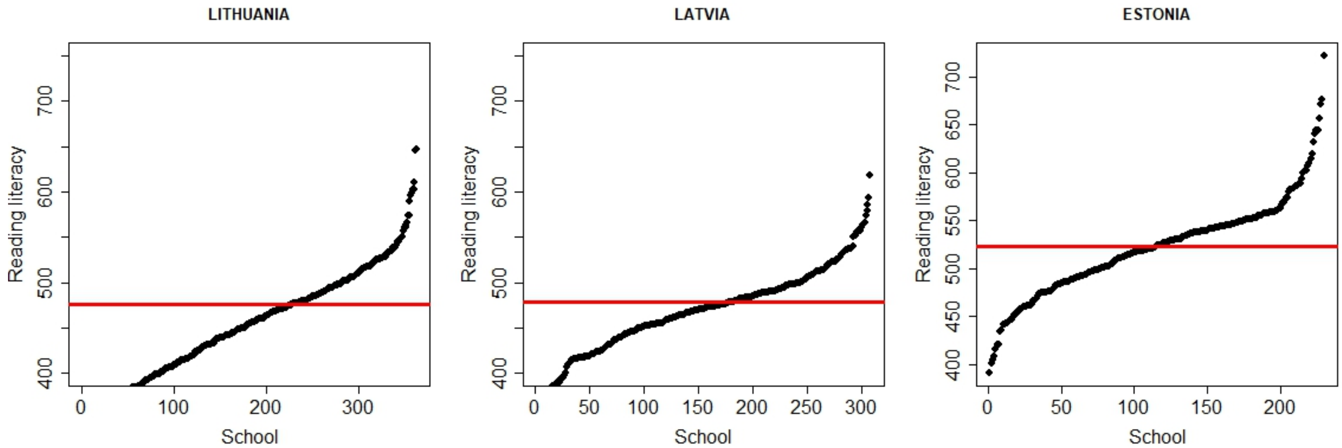

Figure 2: Mean reading literacy scores by school

Table 2 shows differences in the mean of reading literacy between female and male students. Trend towards female students being more successful in reading literacy can be observed.

	Lithuania		Latvia		Estonia	
	Frequency	Mean	Frequency	Mean	Frequency	Mean
Female students	3377	495.63	2685	494.84	2651	538.35
Male students	3508	456.97	2618	462.04	2665	507.7

Table 2: Mean reading literacy distribution among female and male students

Table 3 describes percentages of students at different proficiency levels. PISA scales are divided into 6 levels, where simplest tasks correspond to level 1 (at or above 357.77 in Table 3) and tasks which are the most challenging correspond to level 6 (at or above 669.3 in Table 3) [24]. Table 3 indicates that the distribution of proficiency levels in Lithuania is similar to that of Latvia, while Estonia has a higher proportion of students achieving higher levels of proficiency.

	Lithuania			Latvia			Estonia	
Benchmark	Percentage	SE		Percentage	SE		Percentage	SE
At or above 357.77	88.5	0.53		90.39	0.57		95.88	0.38
At or above 420.07	71.73	0.82		73.29	0.77		86.27	0.68
At or above 482.38	48.7	0.84		49.28	0.89		66.93	1.02
At or above 544.68	24.55	0.62		24.14	0.81		41.25	0.85
At or above 606.99	7.9	0.37		7.75	0.47		18.8	0.7
At or above 669.3	1.33	0.2		1.25	0.2		5.77	0.49

Table 3: Share of students by reading literacy proficiency levels

Additionally, statistics of mean reading literacy in different types of areas and cities can be observed in Table 4. From there it appears that students from schools which are located in rural areas tend to have lower academic achievement compared to those in larger cities.

	Lithuania				
	Freq	Mean	s.e.	SD	s.e
Rural area	1342	427.01	4.66	89.04	2.19
Small town	1488	470.57	3.43	86.67	2.09
Town	1355	477.09	4.13	90.92	2.38
City	2700	504.65	2.85	91.35	1.78
	Latvia				
Rural area	756	445.94	4.64	84.46	2.15
Small town	1305	466.63	3.43	83.22	2.32
Town	1528	481.73	2.76	89.75	1.99
City	1551	497.54	3.06	89.93	1.98
No response	163	528.28	9.31	88.68	5.71
		Estonia			
Rural area	930	517.34	5.02	95.9	2.39
Small town	1210	515.7	3.03	86.75	1.94
Town	1532	517.46	3.22	91.5	2.21
City	1644	536.82	3.43	95.12	1.72

Table 4: Mean achievement score by area type in which school is located

Finally, missing value analysis was conducted. It determined that the most appropriate and straightforward method for handling these missing values was list-wise deletion since this did not results in a significant loss of data. After implementing list-wise deletion, we were left with the final datasets that were used for the analysis. Final Lithuanian sample contains 357 schools with 5618 observations, Latvian sample - 289 schools with 4158 observations and Estonian sample - 229 schools with 4638 observations.

5.2 Modelling self-regulated learning and reading literacy in Lithuania

Table 5 shows the results of a two-level hierarchical linear model (HLM) that was used to examine the relationship between reading literacy and components of self-regulated learning in Lithuania. The model includes both school- and student-level variables as predictors, and reading achievement as the dependent or outcome variable. Student-level predictors include components of self-regulated learning while school-level predictors include school type, area in which school is located and average school's economic, social and cultural status.

Model 1 in table 5, or unconditional model, estimated variance at both student and school levels. The estimate of the intercept $\hat{\gamma}_{00}$ is equal to 455.995 and indicates the mean reading achievement for the overall sample. Student-level variance estimate $\hat{\sigma}^2$ is equal to 5860.445 and school-level variance estimate $\hat{\tau}_{00}$ is 2272.123. Interclass correlation coefficient (ICC) is equal to 0.28 which indicates that around 28% of difference in student results can be attributed to school-level variation. The estimates of the unconditional model suggest that there may be factors at the student and school levels that are contributing to the variance in reading literacy, and therefore, the hierarchical structure of the data should be taken into account. Hence, hierarchical linear modelling approach is applicable in this case.

Unconditional model in table 5 was expanded by including predictors from student-level (*Model 2* in table 5) which are related to self-regulated learning, specifically regulation of cognition and motivation. Addition of these predictors improved ICC value from 0.28 to 0.23 however variance coming from differences between schools remained relatively high. Hence, combined model (*Model 3* in table 5) with school-level predictors was build. Additional predictors reduced variance unexplained by school-level factors to 10% (*ICC=0.1*). Improved model fit is also demonstrated by reduced AIC and BIC values.

Effects of different predictors were estimated as well. Components of cognition or usage of specific learning strategies had a strongest positive effect on reading achievement from all self-regulated learning components for Lithuanian students. Out of two strategies, *METASUM* had a stronger significant effect with $p < .001$. An increase of one unit in the index of usefulness of using summarizing as a reading strategy will result in a 17.458 point increase in a student's reading score. Hence, hypothesis number 1 is confirmed.

Out of 7 indexes related to student motivation included in the model, 4 were found to be statistically significant predictors. The index of perception of competence in reading tasks (*SCREADCOMP*) had a strongest positive relationship, with effect size of 16.355 and $p < .01$. As expected, one unit increase in the index of perceived difficulty in reading (*SCREADDIFF*) will result in reading score reduced by 9.131 with $p < .001$. However, index of fear of failure (*GFOFAIL*) had, although not very strong, but a positive effect of 1.944, with $p < .05$. Enjoyment of reading (*JOYREAD*) was found to have a positive effect on reading score as well with $p < .001$, however only by 4 points. Additionally, controlling for the gender variable (*MALE*) revealed a negative effect on reading literacy score with $p < .01$. Being a male student reduces reading score by 8.7. Student's index of resilience (*RESILIENCE*), motivation to master tasks (*WORKMAST*) and goal setting (*MASTGOAL*) were rejected as having no significant relationship with reading achievement with $p > .05$. In this case, hypotheses number 2, 3 and 4 were confirmed only partially since not all variables related to self-efficacy and motivation were significant or had a desired effect.

The only significant school-level variable was average school's economic, social and cultural status (*SCHOOL_ESCS*), with $p < .001$. It had the strongest effect size from all predictors. One unit increase in school's ESCS leads to 61 point improvement in reading score.

Further subchapter compares results described above with the analogous models build for Latvia and Estonia.

	Lithuania		
	Model 1	Model 2	Model 3
Intercept	455.995*** (3.35)	468.210*** (3.28)	476.8*** (8.58)
MALE	-	-7.279** (2.76)	-8.7** (2.74)
Regulation of cognition			
UNDREM	-	11.717*** (1.39)	11.046*** (1.39)
METASUM	-	18.097*** (1.39)	17.458*** (1.38)
Regulation of motivation			
JOYREAD	-	4.836*** (1.22)	4.053*** (1.20)
RESILIENCE	-	2.187 (1.74)	2.119 (1.70)
SCREADCOMP	-	16.485*** (1.43)	16.355** (1.42)
SCREADDIFF	-	-9.172*** (1.17)	-9.131*** (1.17)
GFOFAIL	-	1.940* (0.97)	1.944* (0.96)
WORKMAST	-	-2.896 (1.51)	-2.481 (1.48)
MASTGOAL	-	0.523 (1.40)	0.708 (1.41)
School level variables			
WHICH_AREA	-	-	0.99 (2.09)
PUBLIC	-	-	10.708 (11.23)
SCHOOL_ESCS	-	-	61.629*** (6.59)
Level 1 variance	5860.449	4587.414	4549.864
Level 2 variance	2272.123	1394.673	513.4475
ICC	0.28	0.23	0.1
N (schools)	357	357	357
Observations	5618	5618	5618
AIC	65458.41	64008.86	63670.62
BIC	65471.67	64081.83	63763.49

Notes: Standard errors are in parentheses. * $p<.05$. **$p<.01$. *** $p<.001$.

Table 5: Results from two-level hierarchical linear model on Lithuania's reading achievement in PISA 2018.

5.3 Model comparison with Latvia and Estonia

Table 6 summarizes the results from analogous two-level hierarchical linear model performed for Latvia and Estonia. In both cases, the intercepts from the unconditional model are higher compared to Lithuania's (see *Model 1* in table 6 for Latvia and Estonia). This is probably because Latvia and Estonia had a higher mean score for reading achievement in the 2018 assessment overall. Estimated variance at school-level is much lower than in Lithuania. 12% and 15% of variance is attributed to schools in Latvia and Estonia respectively (see *Model 1* in table 6 for Latvia and Estonia) compared to 28% in Lithuania. Hence, the majority of variance is explained by the student related factors within schools. When controlling for school level predictors, ICC values reduced to very small - 0.04 and 0.07 respectively (see *Model 3* in table 6 for Latvia and Estonia).

The results from the combined model in some cases differed from those of the model built for Lithuania. Motivational components of self-regulated learning, such as enjoyment of reading (*JOYREAD*), perceived competence (*SCREADCOMP*), and fear of failure (*GFOFAIL*), appear to be stronger predictors of reading literacy in both countries, with $p < .01$. In Latvia's case, one unit change in index for the enjoyment of reading leads to 11.955 point increase in reading score. In Estonia, this effect size is even bigger and equal to 13.256. Index of competence perception in Latvia has a strongest positive effect from all three countries. In comparison with Lithuania and Latvia, index of fear of failure has a strongest positive effect in Estonia. One point shift in this index leads to reading score increasing by 6.889. Additionally, unlike in Lithuania, gender has no significant effect on reading score. And in a similar manner, student's index of resilience (*RESILIENCE*), motivation to master tasks (*WORKMAST*) and goal setting (*MASTGOAL*) were rejected as having no significant relationship with reading achievement Latvia and Estonia as well, with $p > .05$.

As in Lithuania's case, the only significant school-level variable was found to be average school's social, economic and cultural status (*SCHOOL_ESCS*). However, the effect sizes of this predictor in both countries were much lower compared to Lithuania's: 35.218 in Latvia and 47.773 in Estonia.

	Latvia			Estonia		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	476.85*** (2.844)	488.754*** (2.53)	487.526*** (6.08)	528.565*** (3.63)	523.237*** (3.29)	532.546*** (7.24)
MALE	-	-4.558 (3.43)	-5.444 (3.40)	-	4.120 (4.95)	3.512 (4.94)
Regulation of cognition						
UNDREM	-	14.007*** (1.20)	13.451*** (1.22)	-	12.989** (1.37)	12.824*** (1.36)
METASUM	-	18.463*** (1.55)	18.125*** (1.55)	-	19.476** (1.49)	19.143*** (1.51)
Regulation of motivation						
JOYREAD	-	12.964*** (1.71)	11.955*** (1.73)	-	13.587** (1.61)	13.256*** (1.61)
RESILIENCE	-	1.919 (2.50)	1.296 (2.48)	-	2.878 (1.49)	2.586 (1.43)
SCREADCOMP	-	21.7054*** (2.16)	21.48*** (2.17)	-	20.079*** (1.72)	19.928*** (1.70)
SCREADDIFF	-	-10.076*** (1.72)	-9.961*** (1.73)	-	-9.499*** (1.62)	-9.314*** (1.61)
GFOFAIL	-	4.732** (1.60)	4.545** (1.59)	-	7.219*** (1.40)	6.889*** (1.38)
WORKMAST	-	3.037 (1.58)	3.12* (1.58)	-	-1.372 (1.59)	-1.309 (1.58)
MASTGOAL	-	-3.748* (1.73)	-3.288 (1.70)	-	-0.167 (1.57)	-0.411 (1.56)
School level variables						
WHICH_AREA	-	-	1.948 (1.81)	-	-	-4.228* (2.06)
PUBLIC	-	-	-20.549 (29.13)	-	-	-13.485 (12.54)
SCHOOL_ESCS	-	-	35.218*** (5.77)	-	-	47.773*** (7.03)
Level 1 variance	6429.012	4416.22	4414.822	6679.51	4804.122	4791.932
Level 2 variance	885.929	437.2074	203.5267	1184.61	630.479	369.538
ICC	0.12	0.09	0.04	0.15	0.11	0.07
N (schools)	289	289	289	229	229	229
Observations	4158	4158	4158	4638	4638	4638
AIC	48596.11	46984.59	46871.58	54365.92	52801.94	52716.52
BIC	48608.78	47054.25	46960.24	54378.8	52872.81	52806.71

Notes: Standard errors are in parentheses. * $p<.05$. **$p<.01$. *** $p<.001$.

Table 6: Results from two-level hierarchical linear model in Latvia and Estonia

6 Conclusions

The main goals of this thesis were to investigate the Programme of International Student Assessment (PISA) 2018 data from Lithuania, apply self-regulated learning theory to the data, and evaluate the effect and significance of self-regulated learning skills on reading literacy in Lithuania. This involved building a two-level hierarchical linear model and comparing the results obtained in Lithuania with those from Latvia and Estonia. The results of this work can be summarised as follows:

1. Exploratory analysis of the PISA 2018 revealed that Lithuania had the lowest mean reading score between three Baltic countries with Latvia being a close second. Estonia is an outlier with being on a higher level on a proficiency scale. Closer look to gender differences reveal that female students

has a higher average reading score in all three countries. Another common similarity is that students who attend school in a bigger cities tend to score higher on average than those living in smaller towns or rural areas. Initial exploratory analysis indicates Lithuania and Latvia being more similar to each other in terms of achievement levels and characteristics.

2. In Lithuania, learning strategies were found to be significant predictors of reading literacy. This supported theoretical assumptions [30, 29, 4] as well as findings from Shanghai [36], Greece [14] and Turkey [11]. These finding were also supported by models build for Latvia and Estonia, which confirmed that use of metacognitive strategies is important in attaining a higher score in reading.

3. Feeling competent was found to be a significant factor in success in achievement in Lithuania. This supports previous assumptions made by Pintrich and Zimmerman [30, 29, 4] and findings from research conducted in Austria [10]. Additionally, feeling competent in reading was a stronger predictor of success in Latvia and Estonia compared to Lithuania. However, self-efficacy was not found to be a significant factor in Lithuania, Latvia, or Estonia, which contradicts previous research findings in United States as well as theoretical assumptions [30, 29, 4].

4. Enjoyment of reading was found to be a significant predictor of reading literacy in Lithuania. This finding supports previous research [30, 36, 11]. Furthermore, the effect of enjoyment of reading on literacy was stronger in Latvia and Estonia compared to Lithuania. However, in contrast to previous findings from China [31] and Greece [14], motivation to master tasks and goal orientation were not found to be significant predictors of reading literacy in Lithuania or the other Baltic countries, which is a surprising result. Overall, enjoyment of reading appears to be an important factor in predicting reading literacy in the Baltic region.

5. The results on the effects of negative emotions on reading performance confirm that the relationship is not straightforward [30]. As expected, the perception of difficulty had a negative impact on reading literacy scores in Lithuania, and a similar effect was observed in Latvia and Estonia. However, fear of failure was found to have a positive impact, stronger in Latvia and Estonia. This suggests that certain negative emotions may make students more determined and focused on succeeding.

6. The economic, social, and cultural status of a school had a stronger impact on student performance in Lithuania compared to the other two countries. This is also confirmed by comparing characteristics of all three models, where unexplained variance on a school-level (Table 5, *Model 1*) was much higher in Lithuania than in other two countries (Table 6).

7. To summarize, while some self-regulated learning components had an effect on reading literacy for Lithuanian students in PISA 2018, overall, Lithuanian students appear to be weaker in terms of self-regulated learning compared to Latvian and Estonian students, as evidenced by the smaller effect sizes of these components in Lithuania. This finding supports the results of a study conducted after the first quarantine in Lithuania, which found that Lithuanian students struggled with certain self-regulation skills while studying remotely [33]. Therefore, the results of this analysis may possibly prompt further research into the specific factors that contribute to these differences between Baltic countries and why Lithuanian students tend to be weaker at self-regulation, as well as potential ways to improve it.

References

[1] A. Bandura, "Social foundations of thought and action." Englewood Cliffs, NJ, 1986, 23-28, 1986.

[2] A. Karakolidis, V. Pitsia, J. Cosgrove, "Multilevel Modelling of International Large-Scale Assessment Data." *Methodology for Multilevel Modeling in Educational Research*, Springer, Singapore, 2022, 141-159.

[3] B. T. Prasetyanto, N. A. Drajati, and E. Indriyani Putri Handayani. "Dwelling Into the Realm of Self-Regulated Learning within Students Stories in Extensive Reading Activities." *EDULANGUE*, 4(1), 2021, 107-139.

[4] B. Zimmerman, "Attaining self-regulation: A social cognitive perspective." *Handbook of self-regulation*, Academic press, 2000, 13-39.

[5] B. Zimmerman, and D. H. Schunk, eds. Self-regulated learning and academic achievement: Theoretical perspectives. Routledge, 2001.

[6] C. Dignath, and M. VJ Veenman. "The role of direct strategy instruction and indirect activation of self-regulated learning—Evidence from classroom observation studies." *Educational Psychology Review*, 33(2), 2021, 489-533.

[7] C. Dignath, G. Büttner. "Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level." *Metacognition and learning*, 3(3), 2008, 231-264.

[8] D. J. Rivers, M. Nakamura, M. Vallance. "Online self-regulated learning and achievement in the era of change." *Journal of Educational Computing Research*, 60(1), 2022, 104-131.

[9] D. Taranto, and M. T. Buchanan. "Sustaining lifelong learning: A self-regulated learning (SRL) approach." *Discourse and Communication for Sustainable Education*, 11(1), 2020, 5-15.

[10] E. R. Pelikan, et al. "Learning during COVID-19: the role of self-regulated learning, motivation, and procrastination for perceived competence." *Zeitschrift für Erziehungswissenschaft*, 24(2), 2021, 393-418.

[11] E. Tavsancil, O. Yildirim, and S. BILICAN DEMIR. "Direct and indirect effects of learning strategies and reading enjoyment on PISA 2009 reading performance." *Eurasian Journal of Educational Research*, 19(82), 2019, 169-190.

[12] E. Y. Kogar, "An Investigation of the Mediating Role of Various Variables in the Effect of Both Gender and Economic, Social and Cultural Status on Reading Literacy." *International Journal of Progressive Education*, 17(1), 2021, 376-391.

[13] H. Woltman, et al, "An introduction to hierarchical linear modeling." *Tutorials in quantitative methods for psychology*, 8(1), 2012, 52-69.

[14] I. Katsantonis, "Self-regulated learning and reading comprehension: The effects of gender, motivation and metacognition." *Hellenic Journal of Psychology*, 17(3), 2020, 286-307.

[15] I. Koyuncu, and T. Fırat. "Investigating reading literacy in PISA 2018 assessment." *International Electronic Journal of Elementary Education*, 13(2), 2020, 263-275.

[16] J. Jerrim, et al. "To weight or not to weight?: the case of PISA data." *Proceedings of the XXVI Meeting of the Economics of Education Association*, Murcia, Spain, 2017.

[17] J. Lorah, "Analyzing Large-Scale Assessment Data with Multilevel Analyses: Demonstration Using the Programme for International Student Assessment (PISA) 2018 Data." *Methodology for Multilevel Modeling in Educational Research*, Springer, Singapore, 2022, 121-139.

[18] J. Mang, et al, "Sampling weights in multilevel modelling: an investigation using PISA sampling structures." *Large-Scale Assessments in Education* 9(1), 2021, 1-39.

[19] J. Mikk, "Explaining the difference between PISA 2009 reading scores in Finland and Estonia." *Educational Research and Evaluation*, 21(4), 2015, 324-342.

[20] K. Lau, E. Sui-chu Ho, "Reading performance and self-regulated learning of Hong Kong students: What we learnt from PISA 2009." *The Asia-Pacific Education Researcher*, 25(1), 2016, 159-171.

[21] M. Boekaerts, "Motivated learning: Studying student situation transactional units." *European journal of psychology of education*, 14(1), 1999, 41-55.

[22] OECD (2009), PISA Data Analysis Manual: SPSS, Second Edition, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264056275-en.

[23] OECD (2019), PISA 2018 Assessment and Analytical Framework, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b25efab8-en.

[24] OECD (2019), PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en.

[25] OECD (2019), PISA 2018 Results (Volume II): Where All Students Can Succeed, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en.

[26] OECD (2019), PISA 2018 Technical Report, PISA, OECD Publishing https://www.oecd.org/pisa/data/pisa2018technicalreport/

[27] P. Bailey, et al. "Package 'WeMix'.", 2020.

[28] P. Linnakyla, A. Malin, and K. Taube. "Factors behind low reading literacy achievement." *Scandinavian journal of educational research*, 48(3), 2004, 231-249.

[29] P. R. Pintrich, "The role of goal orientation in self-regulated learning." *Handbook of self-regulation.* Academic Press, 2000, 451-502.

[30] P. R. Pintrich, and E. V. De Groot. "Motivational and self-regulated learning components of classroom academic performance." *Journal of educational psychology*, 82(1), 1990, 33.

[31] Q. Qian, K. Lau, "The effects of achievement goals and perceived reading instruction on Chinese student reading performance: Evidence from PISA 2018." *Journal of Research in Reading*, 45(1), 2022, 137-156.

[32] R. Hori, and M. Fujii. "Impact of using ICT for learning purposes on self-efficacy and persistence: Evidence from Pisa 2018." *Sustainability* 13(11), 2021, 6463.

[33] R. Jusienė, et al. "Nuotolinis vaikų ugdymas pandemijos dėl COVID-19 metu: grėsmės ir galimybės ekosisteminiu požiūriu." 2021.

[34] S. Fazriah, D. Irwandi, and D. Fairusi. "Relationship of self-regulated learning with student learning outcomes in chemistry study." *Journal of Physics: Conference Series*. 1836(1), IOP Publishing, 2021.

[35] V. E. Lee, "Using hierarchical linear modeling to study social contexts: The case of school effects." *Educational psychologist*, 35(2), 2000, 125-141.

[36] X. Qi, "Effects of self-regulated learning on student's reading literacy: Evidence from Shanghai." *Frontiers in Psychology*, 11, 2021, 555849.

7 Appendix A: R script

In the script printed here only part for combined model is present. Unconditional and only student-level models were following similar structure. Full script, with unconditional and student-level model as well as exploratory analysis can be hound here: https://github.com/k4rinaa/MasterThesis/blob/main/KarinaSil and in a separate file attached.

```
library("intsvy")
library("WeMix")

## Data import and adjustment of certain variables, their naming
pisa2018_all <- pisa.select.merge(folder = file.path(getwd(), "PISA 2018"),
                           school.file = "CY07_MSU_SCH_QQQ.sav",
                           student.file = "CY07_MSU_STU_QQQ.sav",
                           student = c("CNT",
                                       "CNTSCHID",
                                       "STRATUM",
                                       "ST004D01T",
                                       "ESCS",
                                       "UNDREM",
                                       "METASUM",
                                       "JOYREAD",
                                       "SCREADCOMP",
                                       "SCREADDIFF",
                                       "WORKMAST",
                                       "GFOFAIL",
                                       "RESILIENCE",
                                       "MASTGOAL"),
                           school = c("CNTSCHID",
                                      "SC001Q01TA",
                                      "SC013Q01TA",
                                      "SCHSIZE",
                                      "W_SCHGRNRABWT"),
                           countries = c("LTU", "LVA", "EST"))

# Which area school is located in (town, small town, village, etc.)
pisa2018_all$WHICH_AREA <- pisa2018_all$SC001Q01TA

# Public or private school adjusted
pisa2018_all$PUBLIC <- NA
pisa2018_all$PUBLIC[pisa2018_all$SC013Q01TA==1] <- 0
pisa2018_all$PUBLIC[pisa2018_all$SC013Q01TA==2] <- 1

# Adjusting Gender variable (1=female, 2=male) (converting to binary) - MALE
pisa2018_all$MALE <- NA
pisa2018_all$MALE[pisa2018_all$ST004D01T==1] <- 0
pisa2018_all$MALE[pisa2018_all$ST004D01T==2] <- 1
```

```
## Separate datasets for each country

# Lithuania
pisa2018 <- pisa2018_all[pisa2018_all$CNT=='LTU',]
# Latvia
pisa2018_LVA <- pisa2018_all[pisa2018_all$CNT=='LVA',]
# Estonia
pisa2018_EST <- pisa2018_all[pisa2018_all$CNT=='EST', !grepl("GLCM", names(pisa2018_all))]

## Calculation of cluster weights
# CLUSTER WEIGHT - LTU
# Student sample size in school j
pisa2018 <- pisa2018 %>%
  group_by(CNTSCHID) %>%
  mutate(count = n())

# Sum of total student weights in school j
pisa2018 <- pisa2018 %>%
  group_by(CNTSCHID) %>%
  mutate(w_st_total_j = sum(W_FSTUWT))

# Cluster weight for student i in school j:
cluster_weight <- (pisa2018$count/pisa2018$w_st_total_j)*pisa2018$W_FSTUWT
pisa2018 <- cbind(pisa2018, cluster_weights=cluster_weight)
pisa2018 <- as.data.frame(pisa2018)

#--------------------------------#

# CLUSTER WEIGHT - LVA
# Student sample size in school j
pisa2018_LVA <- pisa2018_LVA %>%
  group_by(CNTSCHID) %>%
  mutate(count = n())

# Sum of total student weights in school j
pisa2018_LVA <- pisa2018_LVA %>%
  group_by(CNTSCHID) %>%
  mutate(w_st_total_j = sum(W_FSTUWT))

# Cluster weight for student i in school j:
cluster_weight <- (pisa2018_LVA$count/pisa2018_LVA$w_st_total_j)*pisa2018_LVA$W_FSTUWT
pisa2018_LVA <- cbind(pisa2018_LVA, cluster_weights=cluster_weight)
pisa2018_LVA <- as.data.frame(pisa2018_LVA)

#--------------------------------#
```

```
# CLUSTER WEIGHT - EST
# Student sample size in school j
pisa2018_EST <- pisa2018_EST %>%
  group_by(CNTSCHID) %>%
  mutate(count = n())

# Sum of total student weights in school j
pisa2018_EST <- pisa2018_EST %>%
  group_by(CNTSCHID) %>%
  mutate(w_st_total_j = sum(W_FSTUWT))

# Cluster weight for student i in school j:
cluster_weight <- (pisa2018_EST$count/pisa2018_EST$w_st_total_j)*pisa2018_EST$W_FSTUWT
pisa2018_EST <- cbind(pisa2018_EST, cluster_weights=cluster_weight)
pisa2018_EST <- as.data.frame(pisa2018_EST)

## Missing values analysis

# How many missing values?
sum(is.na(pisa2018))
sum(is.na(pisa2018_LVA))
sum(is.na(pisa2018_EST))

# How many missing values by column?
colSums(is.na(pisa2018))
colSums(is.na(pisa2018_LVA))
colSums(is.na(pisa2018_EST))

# Which columns have missing values?
names(which(colSums(is.na(pisa2018))>0))
names(which(colSums(is.na(pisa2018_LVA))>0))
names(which(colSums(is.na(pisa2018_EST))>0))

## Handling missing values
# Omitting NA values by listwise deletion
pisa2018_2 <- na.omit(pisa2018)
pisa2018_2_LVA <- na.omit(pisa2018_LVA)
pisa2018_2_EST <- na.omit(pisa2018_EST)

#--------------------------------#

# Create variable for average school ESCS

pisa2018final <- pisa2018final %>%
  group_by(CNTSCHID) %>%
  mutate(SCHOOL_ESCS = mean(ESCS))
```

```
pisa2018final_LVA <- pisa2018final_LVA %>%
  group_by(CNTSCHID) %>%
  mutate(SCHOOL_ESCS = mean(ESCS))

pisa2018final_EST <- pisa2018final_EST %>%
  group_by(CNTSCHID) %>%
  mutate(SCHOOL_ESCS = mean(ESCS))

### Modelling part

##### Final combined model
########## Lithuania #######
model2_pv1 <- mix(PV1READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv2 <- mix(PV2READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv3 <- mix(PV3READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv4 <- mix(PV4READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv5 <- mix(PV5READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv6 <- mix(PV6READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv7 <- mix(PV7READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv8 <- mix(PV8READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final)
model2_pv9 <- mix(PV9READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
```

```
                 data = pisa2018final)
model2_pv10 <- mix(PV10READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                  weights = c("cluster_weights", "W_SCHGRNRABWT"),
                  data = pisa2018final)



## COEFFICIENTS
# Averages from all 10 PVs
model2_coef <- rowMeans(cbind(model2_pv1$coef, model2_pv2$coef, model2_pv3$coef,
    model2_pv4$coef, model2_pv5$coef,
                            model2_pv6$coef, model2_pv7$coef, model2_pv8$coef,
                                model2_pv9$coef, model2_pv10$coef))
round(as.data.frame(model2_coef), 3)

## SE
M <- 10

se_coef1 <- mean(model2_pv1$SE[1], model2_pv2$SE[1], model2_pv3$SE[1], model2_pv4$SE[1],
    model2_pv5$SE[1],
                model2_pv6$SE[1], model2_pv7$SE[1], model2_pv8$SE[1], model2_pv9$SE[1],
                    model2_pv10$SE[1])+(1+(1/M))*var(
                  c(model2_pv1$coef[1], model2_pv2$coef[1], model2_pv3$coef[1],
                      model2_pv4$coef[1],
                    model2_pv5$coef[1], model2_pv6$coef[1], model2_pv7$coef[1],
                        model2_pv8$coef[1],
                    model2_pv9$coef[1], model2_pv10$coef[1]))

se_coef2 <- mean(model2_pv1$SE[2], model2_pv2$SE[2], model2_pv3$SE[2], model2_pv4$SE[2],
    model2_pv5$SE[2],
                model2_pv6$SE[2], model2_pv7$SE[2], model2_pv8$SE[2], model2_pv9$SE[2],
                    model2_pv10$SE[2])+(1+(1/M))*var(
                  c(model2_pv1$coef[2], model2_pv2$coef[2], model2_pv3$coef[2],
                      model2_pv4$coef[2],
                    model2_pv5$coef[2], model2_pv6$coef[2], model2_pv7$coef[2],
                        model2_pv8$coef[2],
                    model2_pv9$coef[2], model2_pv10$coef[2]))

se_coef3 <- mean(model2_pv1$SE[3], model2_pv2$SE[3], model2_pv3$SE[3], model2_pv4$SE[3],
    model2_pv5$SE[3],
                model2_pv6$SE[3], model2_pv7$SE[3], model2_pv8$SE[3], model2_pv9$SE[3],
                    model2_pv10$SE[3])+(1+(1/M))*var(
                  c(model2_pv1$coef[3], model2_pv2$coef[3], model2_pv3$coef[3],
                      model2_pv4$coef[3],
                    model2_pv5$coef[3], model2_pv6$coef[3], model2_pv7$coef[3],
                        model2_pv8$coef[3],
```

```
                    model2_pv9$coef[3], model2_pv10$coef[3]))

se_coef4 <- mean(model2_pv1$SE[4], model2_pv2$SE[4], model2_pv3$SE[4], model2_pv4$SE[4],
    model2_pv5$SE[4],
                model2_pv6$SE[4], model2_pv7$SE[4], model2_pv8$SE[4], model2_pv9$SE[4],
                    model2_pv10$SE[4])+(1+(1/M))*var(
                  c(model2_pv1$coef[4], model2_pv2$coef[4], model2_pv3$coef[4],
                      model2_pv4$coef[4],
                    model2_pv5$coef[4], model2_pv6$coef[4], model2_pv7$coef[4],
                        model2_pv8$coef[4],
                    model2_pv9$coef[4], model2_pv10$coef[4]))

se_coef5 <- mean(model2_pv1$SE[5], model2_pv2$SE[5], model2_pv3$SE[5], model2_pv4$SE[5],
    model2_pv5$SE[5],
                model2_pv6$SE[5], model2_pv7$SE[5], model2_pv8$SE[5], model2_pv9$SE[5],
                    model2_pv10$SE[5])+(1+(1/M))*var(
                  c(model2_pv1$coef[5], model2_pv2$coef[5], model2_pv3$coef[5],
                      model2_pv4$coef[5],
                    model2_pv5$coef[5], model2_pv6$coef[5], model2_pv7$coef[5],
                        model2_pv8$coef[5],
                    model2_pv9$coef[5], model2_pv10$coef[5]))

se_coef6 <- mean(model2_pv1$SE[6], model2_pv2$SE[6], model2_pv3$SE[6], model2_pv4$SE[6],
    model2_pv5$SE[6],
                model2_pv6$SE[6], model2_pv7$SE[6], model2_pv8$SE[6], model2_pv9$SE[6],
                    model2_pv10$SE[6])+(1+(1/M))*var(
                  c(model2_pv1$coef[6], model2_pv2$coef[6], model2_pv3$coef[6],
                      model2_pv4$coef[6],
                    model2_pv5$coef[6], model2_pv6$coef[6], model2_pv7$coef[6],
                        model2_pv8$coef[6],
                    model2_pv9$coef[6], model2_pv10$coef[6]))

se_coef7 <- mean(model2_pv1$SE[7], model2_pv2$SE[7], model2_pv3$SE[7], model2_pv4$SE[7],
    model2_pv5$SE[7],
                model2_pv6$SE[7], model2_pv7$SE[7], model2_pv8$SE[7], model2_pv9$SE[7],
                    model2_pv10$SE[7])+(1+(1/M))*var(
                  c(model2_pv1$coef[7], model2_pv2$coef[7], model2_pv3$coef[7],
                      model2_pv4$coef[7],
                    model2_pv5$coef[7], model2_pv6$coef[7], model2_pv7$coef[7],
                        model2_pv8$coef[7],
                    model2_pv9$coef[7], model2_pv10$coef[7]))

se_coef8 <- mean(model2_pv1$SE[8], model2_pv2$SE[8], model2_pv3$SE[8], model2_pv4$SE[8],
    model2_pv5$SE[8],
                model2_pv6$SE[8], model2_pv7$SE[8], model2_pv8$SE[8], model2_pv9$SE[8],
                    model2_pv10$SE[8])+(1+(1/M))*var(
```

```
                  c(model2_pv1$coef[8], model2_pv2$coef[8], model2_pv3$coef[8],
                      model2_pv4$coef[8],
                    model2_pv5$coef[8], model2_pv6$coef[8], model2_pv7$coef[8],
                        model2_pv8$coef[8],
                    model2_pv9$coef[8], model2_pv10$coef[8]))

se_coef9 <- mean(model2_pv1$SE[9], model2_pv2$SE[9], model2_pv3$SE[9], model2_pv4$SE[9],
    model2_pv5$SE[9],
                model2_pv6$SE[9], model2_pv7$SE[9], model2_pv8$SE[9], model2_pv9$SE[9],
                    model2_pv10$SE[9])+(1+(1/M))*var(
                  c(model2_pv1$coef[9], model2_pv2$coef[9], model2_pv3$coef[9],
                      model2_pv4$coef[9],
                    model2_pv5$coef[9], model2_pv6$coef[9], model2_pv7$coef[9],
                        model2_pv8$coef[9],
                    model2_pv9$coef[9], model2_pv10$coef[9]))

se_coef10 <- mean(model2_pv1$SE[10], model2_pv2$SE[10], model2_pv3$SE[10], model2_pv4$SE
    [10], model2_pv5$SE[10],
                 model2_pv6$SE[10], model2_pv7$SE[10], model2_pv8$SE[10], model2_pv9$SE[10],
                     model2_pv10$SE[10])+(1+(1/M))*var(
                   c(model2_pv1$coef[10], model2_pv2$coef[10], model2_pv3$coef[10],
                       model2_pv4$coef[10],
                     model2_pv5$coef[10], model2_pv6$coef[10], model2_pv7$coef[10],
                         model2_pv8$coef[10],
                     model2_pv9$coef[10], model2_pv10$coef[10]))

se_coef11 <- mean(model2_pv1$SE[11], model2_pv2$SE[11], model2_pv3$SE[11], model2_pv4$SE
    [11], model2_pv5$SE[11],
                 model2_pv6$SE[11], model2_pv7$SE[11], model2_pv8$SE[11], model2_pv9$SE[11],
                     model2_pv10$SE[11])+(1+(1/M))*var(
                   c(model2_pv1$coef[11], model2_pv2$coef[11], model2_pv3$coef[11],
                       model2_pv4$coef[11],
                     model2_pv5$coef[11], model2_pv6$coef[11], model2_pv7$coef[11],
                         model2_pv8$coef[11],
                     model2_pv9$coef[11], model2_pv10$coef[11]))

se_coef12 <- mean(model2_pv1$SE[12], model2_pv2$SE[12], model2_pv3$SE[12], model2_pv4$SE
    [12], model2_pv5$SE[12],
                 model2_pv6$SE[12], model2_pv7$SE[12], model2_pv8$SE[12], model2_pv9$SE[12],
                     model2_pv10$SE[12])+(1+(1/M))*var(
                   c(model2_pv1$coef[12], model2_pv2$coef[12], model2_pv3$coef[12],
                       model2_pv4$coef[12],
                     model2_pv5$coef[12], model2_pv6$coef[12], model2_pv7$coef[12],
                         model2_pv8$coef[12],
                     model2_pv9$coef[12], model2_pv10$coef[12]))
```

```
se_coef13 <- mean(model2_pv1$SE[13], model2_pv2$SE[13], model2_pv3$SE[13], model2_pv4$SE
    [13], model2_pv5$SE[13],
                model2_pv6$SE[13], model2_pv7$SE[13], model2_pv8$SE[13], model2_pv9$SE[13],
                    model2_pv10$SE[13])+(1+(1/M))*var(
                  c(model2_pv1$coef[13], model2_pv2$coef[13], model2_pv3$coef[13],
                      model2_pv4$coef[13],
                    model2_pv5$coef[13], model2_pv6$coef[13], model2_pv7$coef[13],
                        model2_pv8$coef[13],
                    model2_pv9$coef[13], model2_pv10$coef[13]))

se_coef14 <- mean(model2_pv1$SE[14], model2_pv2$SE[14], model2_pv3$SE[14], model2_pv4$SE
    [14], model2_pv5$SE[14],
                model2_pv6$SE[14], model2_pv7$SE[14], model2_pv8$SE[14], model2_pv9$SE[14],
                    model2_pv10$SE[14])+(1+(1/M))*var(
                  c(model2_pv1$coef[14], model2_pv2$coef[14], model2_pv3$coef[14],
                      model2_pv4$coef[14],
                    model2_pv5$coef[14], model2_pv6$coef[14], model2_pv7$coef[14],
                        model2_pv8$coef[14],
                    model2_pv9$coef[14], model2_pv10$coef[14]))

model2_se <- rbind(Intercept=se_coef1, MALE=se_coef2, UNDREM=se_coef3, METASUM=se_coef4,
    JOYREAD=se_coef5,
                 RESILIENCE=se_coef6, SCREADCOMP=se_coef7, SCREADDIFF=se_coef8, GFOFAIL=
                     se_coef9,
                 WORKMAST=se_coef10, MASTGOAL=se_coef11, WHICH_AREA=se_coef12, PUBLIC=
                     se_coef13, SCHOOL_ESCS=se_coef14)


## t-values
model2_t <- model2_coef/model2_se

## p-values
model2_p <- round(2 * pnorm( abs(model2_t), lower.tail=FALSE), 3)


## ICC
model2_ICC <- mean(model2_pv1$ICC, model2_pv2$ICC, model2_pv3$ICC, model2_pv4$ICC,
    model2_pv5$ICC,
                 model2_pv6$ICC, model2_pv7$ICC, model2_pv8$ICC, model2_pv9$ICC,
                     model2_pv10$ICC)


# School variance
mean(model2_pv1$vars[1], model2_pv2$vars[1], model2_pv3$vars[1], model2_pv4$vars[1],
    model2_pv5$vars[1],
     model2_pv6$vars[1], model2_pv7$vars[1], model2_pv8$vars[1], model2_pv9$vars[1],
         model2_pv10$vars[1])
```

```
# Student level variance
mean(model2_pv1$vars[2], model2_pv2$vars[2], model2_pv3$vars[2], model2_pv4$vars[2],
    model2_pv5$vars[2],
     model2_pv6$vars[2], model2_pv7$vars[2], model2_pv8$vars[2], model2_pv9$vars[2],
        model2_pv20$vars[2])


## AIC
model2_loglik <- mean(model2_pv1$lnl, model2_pv2$lnl, model2_pv3$lnl, model2_pv4$lnl,
    model2_pv5$lnl,
                    model2_pv6$lnl, model2_pv7$lnl, model2_pv8$lnl, model2_pv9$lnl,
                        model2_pv10$lnl)

model2_AIC <- -2 * model2_loglik + 2 * 14

## BIC
model2_BIC <- log(nrow(pisa2018final)) * 14 - 2 * model2_loglik


#--------------------------------#
######## Latvia ######

model2_pv1_lva <- mix(PV1READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv2_lva <- mix(PV2READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv3_lva <- mix(PV3READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv4_lva <- mix(PV4READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv5_lva <- mix(PV5READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv6_lva <- mix(PV6READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
```

```
model2_pv7_lva <- mix(PV7READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv8_lva <- mix(PV8READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv9_lva <- mix(PV9READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                weights = c("cluster_weights", "W_SCHGRNRABWT"),
                data = pisa2018final_LVA)
model2_pv10_lva <- mix(PV10READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF
    + GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                 weights = c("cluster_weights", "W_SCHGRNRABWT"),
                 data = pisa2018final_LVA)

## COEFFICIENTS
# Averages from all 10 PVs
model2_coef_lva <- rowMeans(cbind(model2_pv1_lva$coef, model2_pv2_lva$coef,
    model2_pv3_lva$coef, model2_pv4_lva$coef, model2_pv5_lva$coef,
                           model2_pv6_lva$coef, model2_pv7_lva$coef, model2_pv8_lva$coef,
                               model2_pv9_lva$coef, model2_pv10_lva$coef))
round(as.data.frame(model2_coef_lva), 3)

## SE
M <- 10

se_coef1_lva <- mean(model2_pv1_lva$SE[1], model2_pv2_lva$SE[1], model2_pv3_lva$SE[1],
    model2_pv4_lva$SE[1], model2_pv5_lva$SE[1],
                   model2_pv6_lva$SE[1], model2_pv7_lva$SE[1], model2_pv8_lva$SE[1],
                       model2_pv9_lva$SE[1], model2_pv10_lva$SE[1])+(1+(1/M))*var(
                     c(model2_pv1_lva$coef[1], model2_pv2_lva$coef[1], model2_pv3_lva$coef
                         [1], model2_pv4_lva$coef[1],
                       model2_pv5_lva$coef[1], model2_pv6_lva$coef[1], model2_pv7_lva$coef
                           [1], model2_pv8_lva$coef[1],
                       model2_pv9_lva$coef[1], model2_pv10_lva$coef[1]))

se_coef2_lva <- mean(model2_pv1_lva$SE[2], model2_pv2_lva$SE[2], model2_pv3_lva$SE[2],
    model2_pv4_lva$SE[2], model2_pv5_lva$SE[2],
                   model2_pv6_lva$SE[2], model2_pv7_lva$SE[2], model2_pv8_lva$SE[2],
                       model2_pv9_lva$SE[2], model2_pv10_lva$SE[2])+(1+(1/M))*var(
                     c(model2_pv1_lva$coef[2], model2_pv2_lva$coef[2], model2_pv3_lva$coef
                         [2], model2_pv4_lva$coef[2],
                       model2_pv5_lva$coef[2], model2_pv6_lva$coef[2], model2_pv7_lva$coef
                           [2], model2_pv8_lva$coef[2],
                       model2_pv9_lva$coef[2], model2_pv10_lva$coef[2]))
```

```
se_coef3_lva <- mean(model2_pv1_lva$SE[3], model2_pv2_lva$SE[3], model2_pv3_lva$SE[3],
    model2_pv4_lva$SE[3], model2_pv5_lva$SE[3],
                    model2_pv6_lva$SE[3], model2_pv7_lva$SE[3], model2_pv8_lva$SE[3],
                        model2_pv9_lva$SE[3], model2_pv10_lva$SE[3])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[3], model2_pv2_lva$coef[3], model2_pv3_lva$coef
                          [3], model2_pv4_lva$coef[3],
                        model2_pv5_lva$coef[3], model2_pv6_lva$coef[3], model2_pv7_lva$coef
                            [3], model2_pv8_lva$coef[3],
                        model2_pv9_lva$coef[3], model2_pv10_lva$coef[3]))

se_coef4_lva <- mean(model2_pv1_lva$SE[4], model2_pv2_lva$SE[4], model2_pv3_lva$SE[4],
    model2_pv4_lva$SE[4], model2_pv5_lva$SE[4],
                    model2_pv6_lva$SE[4], model2_pv7_lva$SE[4], model2_pv8_lva$SE[4],
                        model2_pv9_lva$SE[4], model2_pv10_lva$SE[4])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[4], model2_pv2_lva$coef[4], model2_pv3_lva$coef
                          [4], model2_pv4_lva$coef[4],
                        model2_pv5_lva$coef[4], model2_pv6_lva$coef[4], model2_pv7_lva$coef
                            [4], model2_pv8_lva$coef[4],
                        model2_pv9_lva$coef[4], model2_pv10_lva$coef[4]))

se_coef5_lva <- mean(model2_pv1_lva$SE[5], model2_pv2_lva$SE[5], model2_pv3_lva$SE[5],
    model2_pv4_lva$SE[5], model2_pv5_lva$SE[5],
                    model2_pv6_lva$SE[5], model2_pv7_lva$SE[5], model2_pv8_lva$SE[5],
                        model2_pv9_lva$SE[5], model2_pv10_lva$SE[5])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[5], model2_pv2_lva$coef[5], model2_pv3_lva$coef
                          [5], model2_pv4_lva$coef[5],
                        model2_pv5_lva$coef[5], model2_pv6_lva$coef[5], model2_pv7_lva$coef
                            [5], model2_pv8_lva$coef[5],
                        model2_pv9_lva$coef[5], model2_pv10_lva$coef[5]))

se_coef6_lva <- mean(model2_pv1_lva$SE[6], model2_pv2_lva$SE[6], model2_pv3_lva$SE[6],
    model2_pv4_lva$SE[6], model2_pv5_lva$SE[6],
                    model2_pv6_lva$SE[6], model2_pv7_lva$SE[6], model2_pv8_lva$SE[6],
                        model2_pv9_lva$SE[6], model2_pv10_lva$SE[6])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[6], model2_pv2_lva$coef[6], model2_pv3_lva$coef
                          [6], model2_pv4_lva$coef[6],
                        model2_pv5_lva$coef[6], model2_pv6_lva$coef[6], model2_pv7_lva$coef
                            [6], model2_pv8_lva$coef[6],
                        model2_pv9_lva$coef[6], model2_pv10_lva$coef[6]))

se_coef7_lva <- mean(model2_pv1_lva$SE[7], model2_pv2_lva$SE[7], model2_pv3_lva$SE[7],
    model2_pv4_lva$SE[7], model2_pv5_lva$SE[7],
                    model2_pv6_lva$SE[7], model2_pv7_lva$SE[7], model2_pv8_lva$SE[7],
                        model2_pv9_lva$SE[7], model2_pv10_lva$SE[7])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[7], model2_pv2_lva$coef[7], model2_pv3_lva$coef
                          [7], model2_pv4_lva$coef[7],
```

```
                        model2_pv5_lva$coef[7], model2_pv6_lva$coef[7], model2_pv7_lva$coef
                            [7], model2_pv8_lva$coef[7],
                        model2_pv9_lva$coef[7], model2_pv10_lva$coef[7]))

se_coef8_lva <- mean(model2_pv1_lva$SE[8], model2_pv2_lva$SE[8], model2_pv3_lva$SE[8],
    model2_pv4_lva$SE[8], model2_pv5_lva$SE[8],
                    model2_pv6_lva$SE[8], model2_pv7_lva$SE[8], model2_pv8_lva$SE[8],
                        model2_pv9_lva$SE[8], model2_pv10_lva$SE[8])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[8], model2_pv2_lva$coef[8], model2_pv3_lva$coef
                          [8], model2_pv4_lva$coef[8],
                        model2_pv5_lva$coef[8], model2_pv6_lva$coef[8], model2_pv7_lva$coef
                            [8], model2_pv8_lva$coef[8],
                        model2_pv9_lva$coef[8], model2_pv10_lva$coef[8]))

se_coef9_lva <- mean(model2_pv1_lva$SE[9], model2_pv2_lva$SE[9], model2_pv3_lva$SE[9],
    model2_pv4_lva$SE[9], model2_pv5_lva$SE[9],
                    model2_pv6_lva$SE[9], model2_pv7_lva$SE[9], model2_pv8_lva$SE[9],
                        model2_pv9_lva$SE[9], model2_pv10_lva$SE[9])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[9], model2_pv2_lva$coef[9], model2_pv3_lva$coef
                          [9], model2_pv4_lva$coef[9],
                        model2_pv5_lva$coef[9], model2_pv6_lva$coef[9], model2_pv7_lva$coef
                            [9], model2_pv8_lva$coef[9],
                        model2_pv9_lva$coef[9], model2_pv10_lva$coef[9]))

se_coef10_lva <- mean(model2_pv1_lva$SE[10], model2_pv2_lva$SE[10], model2_pv3_lva$SE[10],
    model2_pv4_lva$SE[10], model2_pv5_lva$SE[10],
                     model2_pv6_lva$SE[10], model2_pv7_lva$SE[10], model2_pv8_lva$SE[10],
                         model2_pv9_lva$SE[10], model2_pv10_lva$SE[10])+(1+(1/M))*var(
                       c(model2_pv1_lva$coef[10], model2_pv2_lva$coef[10], model2_pv3_lva$coef
                           [10], model2_pv4_lva$coef[10],
                         model2_pv5_lva$coef[10], model2_pv6_lva$coef[10], model2_pv7_lva$coef
                             [10], model2_pv8_lva$coef[10],
                         model2_pv9_lva$coef[10], model2_pv10_lva$coef[10]))

se_coef11_lva <- mean(model2_pv1_lva$SE[11], model2_pv2_lva$SE[11], model2_pv3_lva$SE[11],
    model2_pv4_lva$SE[11], model2_pv5_lva$SE[11],
                     model2_pv6_lva$SE[11], model2_pv7_lva$SE[11], model2_pv8_lva$SE[11],
                         model2_pv9_lva$SE[11], model2_pv10_lva$SE[11])+(1+(1/M))*var(
                       c(model2_pv1_lva$coef[11], model2_pv2_lva$coef[11], model2_pv3_lva$coef
                           [11], model2_pv4_lva$coef[11],
                         model2_pv5_lva$coef[11], model2_pv6_lva$coef[11], model2_pv7_lva$coef
                             [11], model2_pv8_lva$coef[11],
                         model2_pv9_lva$coef[11], model2_pv10_lva$coef[11]))

se_coef12_lva <- mean(model2_pv1_lva$SE[12], model2_pv2_lva$SE[12], model2_pv3_lva$SE[12],
    model2_pv4_lva$SE[12], model2_pv5_lva$SE[12],
```

```
                    model2_pv6_lva$SE[12], model2_pv7_lva$SE[12], model2_pv8_lva$SE[12],
                        model2_pv9_lva$SE[12], model2_pv10_lva$SE[12])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[12], model2_pv2_lva$coef[12], model2_pv3_lva$coef
                          [12], model2_pv4_lva$coef[12],
                        model2_pv5_lva$coef[12], model2_pv6_lva$coef[12], model2_pv7_lva$coef
                            [12], model2_pv8_lva$coef[12],
                        model2_pv9_lva$coef[12], model2_pv10_lva$coef[12]))

se_coef13_lva <- mean(model2_pv1_lva$SE[13], model2_pv2_lva$SE[13], model2_pv3_lva$SE[13],
    model2_pv4_lva$SE[13], model2_pv5_lva$SE[13],
                    model2_pv6_lva$SE[13], model2_pv7_lva$SE[13], model2_pv8_lva$SE[13],
                        model2_pv9_lva$SE[13], model2_pv10_lva$SE[13])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[13], model2_pv2_lva$coef[13], model2_pv3_lva$coef
                          [13], model2_pv4_lva$coef[13],
                        model2_pv5_lva$coef[13], model2_pv6_lva$coef[13], model2_pv7_lva$coef
                            [13], model2_pv8_lva$coef[13],
                        model2_pv9_lva$coef[13], model2_pv10_lva$coef[13]))

se_coef14_lva <- mean(model2_pv1_lva$SE[14], model2_pv2_lva$SE[14], model2_pv3_lva$SE[14],
    model2_pv4_lva$SE[14], model2_pv5_lva$SE[14],
                    model2_pv6_lva$SE[14], model2_pv7_lva$SE[14], model2_pv8_lva$SE[14],
                        model2_pv9_lva$SE[14], model2_pv10_lva$SE[14])+(1+(1/M))*var(
                      c(model2_pv1_lva$coef[14], model2_pv2_lva$coef[14], model2_pv3_lva$coef
                          [14], model2_pv4_lva$coef[14],
                        model2_pv5_lva$coef[14], model2_pv6_lva$coef[14], model2_pv7_lva$coef
                            [14], model2_pv8_lva$coef[14],
                        model2_pv9_lva$coef[14], model2_pv10_lva$coef[14]))


model2_se_lva <- rbind(Intercept=se_coef1_lva, MALE=se_coef2_lva, UNDREM=se_coef3_lva,
    METASUM=se_coef4_lva, JOYREAD=se_coef5_lva,
                 RESILIENCE=se_coef6_lva, SCREADCOMP=se_coef7_lva, SCREADDIFF=se_coef8_lva,
                     GFOFAIL=se_coef9_lva,
                 WORKMAST=se_coef10_lva, MASTGOAL=se_coef11_lva, WHICH_AREA=se_coef12_lva,
                     PUBLIC=se_coef13_lva, SCHOOL_ESCS=se_coef14_lva)



## t-values
model2_t_lva <- model2_coef_lva/model2_se_lva

## p-values
model2_p_lva <- round(2 * pnorm( abs(model2_t_lva), lower.tail=FALSE), 3)


## ICC
```

```
model2_ICC_lva <- mean(model2_pv1_lva$ICC, model2_pv2_lva$ICC, model2_pv3_lva$ICC,
    model2_pv4_lva$ICC, model2_pv5_lva$ICC,
                  model2_pv6_lva$ICC, model2_pv7_lva$ICC, model2_pv8_lva$ICC,
                      model2_pv9_lva$ICC, model2_pv10_lva$ICC)


# School variance
mean(model2_pv1_lva$vars[1], model2_pv2_lva$vars[1], model2_pv3_lva$vars[1],
    model2_pv4_lva$vars[1], model2_pv5_lva$vars[1],
     model2_pv6_lva$vars[1], model2_pv7_lva$vars[1], model2_pv8_lva$vars[1],
         model2_pv9_lva$vars[1], model2_pv10_lva$vars[1])

# Student level variance
mean(model2_pv1_lva$vars[2], model2_pv2_lva$vars[2], model2_pv3_lva$vars[2],
    model2_pv4_lva$vars[2], model2_pv5_lva$vars[2],
     model2_pv6_lva$vars[2], model2_pv7_lva$vars[2], model2_pv8_lva$vars[2],
         model2_pv9_lva$vars[2], model2_pv10_lva$vars[2])



## AIC
model2_loglik_lva <- mean(model2_pv1_lva$lnl, model2_pv2_lva$lnl, model2_pv3_lva$lnl,
    model2_pv4_lva$lnl, model2_pv5_lva$lnl,
                     model2_pv6_lva$lnl, model2_pv7_lva$lnl, model2_pv8_lva$lnl,
                         model2_pv9_lva$lnl, model2_pv10_lva$lnl)

model2_AIC_lva <- -2 * model2_loglik_lva + 2 * 14

## BIC
model2_BIC_lva <- log(nrow(pisa2018final_LVA)) * 14 - 2 * model2_loglik_lva

#--------------------------------#
########## Estonia

model2_pv1_est <- mix(PV1READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv2_est <- mix(PV2READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                      weights = c("cluster_weights", "W_SCHGRNRABWT"),
                      data = pisa2018final_EST)
model2_pv3_est <- mix(PV3READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                      weights = c("cluster_weights", "W_SCHGRNRABWT"),
                      data = pisa2018final_EST)
```

```
model2_pv4_est <- mix(PV4READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv5_est <- mix(PV5READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv6_est <- mix(PV6READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv7_est <- mix(PV7READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv8_est <- mix(PV8READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv9_est <- mix(PV9READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF+
    GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                     weights = c("cluster_weights", "W_SCHGRNRABWT"),
                     data = pisa2018final_EST)
model2_pv10_est <- mix(PV10READ~MALE+UNDREM+METASUM+JOYREAD+RESILIENCE+SCREADCOMP+SCREADDIFF
    + GFOFAIL+WORKMAST+MASTGOAL+WHICH_AREA+PUBLIC+SCHOOL_ESCS+(1|CNTSCHID),
                      weights = c("cluster_weights", "W_SCHGRNRABWT"),
                      data = pisa2018final_EST)

names(pisa2018final)

## COEFFICIENTS
# Averages from all 10 PVs

model2_coef_est <- rowMeans(cbind(model2_pv1_est$coef, model2_pv2_est$coef,
    model2_pv3_est$coef, model2_pv4_est$coef, model2_pv5_est$coef,
                               model2_pv6_est$coef, model2_pv7_est$coef, model2_pv8_est$coef
                                   , model2_pv9_est$coef, model2_pv10_est$coef))

round(as.data.frame(model2_coef_est), 3)

## SE
M <- 10

se_coef1_est <- mean(model2_pv1_est$SE[1], model2_pv2_est$SE[1], model2_pv3_est$SE[1],
    model2_pv4_est$SE[1], model2_pv5_est$SE[1],
```

```
                    model2_pv6_est$SE[1], model2_pv7_est$SE[1], model2_pv8_est$SE[1],
                        model2_pv9_est$SE[1], model2_pv10_est$SE[1])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[1], model2_pv2_est$coef[1], model2_pv3_est$coef
                          [1], model2_pv4_est$coef[1],
                        model2_pv5_est$coef[1], model2_pv6_est$coef[1], model2_pv7_est$coef
                            [1], model2_pv8_est$coef[1],
                        model2_pv9_est$coef[1], model2_pv10_est$coef[1]))

se_coef2_est <- mean(model2_pv1_est$SE[2], model2_pv2_est$SE[2], model2_pv3_est$SE[2],
    model2_pv4_est$SE[2], model2_pv5_est$SE[2],
                    model2_pv6_est$SE[2], model2_pv7_est$SE[2], model2_pv8_est$SE[2],
                        model2_pv9_est$SE[2], model2_pv10_est$SE[2])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[2], model2_pv2_est$coef[2], model2_pv3_est$coef
                          [2], model2_pv4_est$coef[2],
                        model2_pv5_est$coef[2], model2_pv6_est$coef[2], model2_pv7_est$coef
                            [2], model2_pv8_est$coef[2],
                        model2_pv9_est$coef[2], model2_pv10_est$coef[2]))

se_coef3_est <- mean(model2_pv1_est$SE[3], model2_pv2_est$SE[3], model2_pv3_est$SE[3],
    model2_pv4_est$SE[3], model2_pv5_est$SE[3],
                    model2_pv6_est$SE[3], model2_pv7_est$SE[3], model2_pv8_est$SE[3],
                        model2_pv9_est$SE[3], model2_pv10_est$SE[3])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[3], model2_pv2_est$coef[3], model2_pv3_est$coef
                          [3], model2_pv4_est$coef[3],
                        model2_pv5_est$coef[3], model2_pv6_est$coef[3], model2_pv7_est$coef
                            [3], model2_pv8_est$coef[3],
                        model2_pv9_est$coef[3], model2_pv10_est$coef[3]))

se_coef4_est <- mean(model2_pv1_est$SE[4], model2_pv2_est$SE[4], model2_pv3_est$SE[4],
    model2_pv4_est$SE[4], model2_pv5_est$SE[4],
                    model2_pv6_est$SE[4], model2_pv7_est$SE[4], model2_pv8_est$SE[4],
                        model2_pv9_est$SE[4], model2_pv10_est$SE[4])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[4], model2_pv2_est$coef[4], model2_pv3_est$coef
                          [4], model2_pv4_est$coef[4],
                        model2_pv5_est$coef[4], model2_pv6_est$coef[4], model2_pv7_est$coef
                            [4], model2_pv8_est$coef[4],
                        model2_pv9_est$coef[4], model2_pv10_est$coef[4]))

se_coef5_est <- mean(model2_pv1_est$SE[5], model2_pv2_est$SE[5], model2_pv3_est$SE[5],
    model2_pv4_est$SE[5], model2_pv5_est$SE[5],
                    model2_pv6_est$SE[5], model2_pv7_est$SE[5], model2_pv8_est$SE[5],
                        model2_pv9_est$SE[5], model2_pv10_est$SE[5])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[5], model2_pv2_est$coef[5], model2_pv3_est$coef
                          [5], model2_pv4_est$coef[5],
                        model2_pv5_est$coef[5], model2_pv6_est$coef[5], model2_pv7_est$coef
                            [5], model2_pv8_est$coef[5],
                        model2_pv9_est$coef[5], model2_pv10_est$coef[5]))
```

```
se_coef6_est <- mean(model2_pv1_est$SE[6], model2_pv2_est$SE[6], model2_pv3_est$SE[6],
    model2_pv4_est$SE[6], model2_pv5_est$SE[6],
                    model2_pv6_est$SE[6], model2_pv7_est$SE[6], model2_pv8_est$SE[6],
                        model2_pv9_est$SE[6], model2_pv10_est$SE[6])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[6], model2_pv2_est$coef[6], model2_pv3_est$coef
                          [6], model2_pv4_est$coef[6],
                        model2_pv5_est$coef[6], model2_pv6_est$coef[6], model2_pv7_est$coef
                            [6], model2_pv8_est$coef[6],
                        model2_pv9_est$coef[6], model2_pv10_est$coef[6]))

se_coef7_est <- mean(model2_pv1_est$SE[7], model2_pv2_est$SE[7], model2_pv3_est$SE[7],
    model2_pv4_est$SE[7], model2_pv5_est$SE[7],
                    model2_pv6_est$SE[7], model2_pv7_est$SE[7], model2_pv8_est$SE[7],
                        model2_pv9_est$SE[7], model2_pv10_est$SE[7])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[7], model2_pv2_est$coef[7], model2_pv3_est$coef
                          [7], model2_pv4_est$coef[7],
                        model2_pv5_est$coef[7], model2_pv6_est$coef[7], model2_pv7_est$coef
                            [7], model2_pv8_est$coef[7],
                        model2_pv9_est$coef[7], model2_pv10_est$coef[7]))

se_coef8_est <- mean(model2_pv1_est$SE[8], model2_pv2_est$SE[8], model2_pv3_est$SE[8],
    model2_pv4_est$SE[8], model2_pv5_est$SE[8],
                    model2_pv6_est$SE[8], model2_pv7_est$SE[8], model2_pv8_est$SE[8],
                        model2_pv9_est$SE[8], model2_pv10_est$SE[8])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[8], model2_pv2_est$coef[8], model2_pv3_est$coef
                          [8], model2_pv4_est$coef[8],
                        model2_pv5_est$coef[8], model2_pv6_est$coef[8], model2_pv7_est$coef
                            [8], model2_pv8_est$coef[8],
                        model2_pv9_est$coef[8], model2_pv10_est$coef[8]))

se_coef9_est <- mean(model2_pv1_est$SE[9], model2_pv2_est$SE[9], model2_pv3_est$SE[9],
    model2_pv4_est$SE[9], model2_pv5_est$SE[9],
                    model2_pv6_est$SE[9], model2_pv7_est$SE[9], model2_pv8_est$SE[9],
                        model2_pv9_est$SE[9], model2_pv10_est$SE[9])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[9], model2_pv2_est$coef[9], model2_pv3_est$coef
                          [9], model2_pv4_est$coef[9],
                        model2_pv5_est$coef[9], model2_pv6_est$coef[9], model2_pv7_est$coef
                            [9], model2_pv8_est$coef[9],
                        model2_pv9_est$coef[9], model2_pv10_est$coef[9]))

se_coef10_est <- mean(model2_pv1_est$SE[10], model2_pv2_est$SE[10], model2_pv3_est$SE[10],
    model2_pv4_est$SE[10], model2_pv5_est$SE[10],
                     model2_pv6_est$SE[10], model2_pv7_est$SE[10], model2_pv8_est$SE[10],
                         model2_pv9_est$SE[10], model2_pv10_est$SE[10])+(1+(1/M))*var(
                       c(model2_pv1_est$coef[10], model2_pv2_est$coef[10], model2_pv3_est$coef
                           [10], model2_pv4_est$coef[10],
```

```
                        model2_pv5_est$coef[10], model2_pv6_est$coef[10], model2_pv7_est$coef
                            [10], model2_pv8_est$coef[10],
                        model2_pv9_est$coef[10], model2_pv10_est$coef[10]))

se_coef11_est <- mean(model2_pv1_est$SE[11], model2_pv2_est$SE[11], model2_pv3_est$SE[11],
    model2_pv4_est$SE[11], model2_pv5_est$SE[11],
                    model2_pv6_est$SE[11], model2_pv7_est$SE[11], model2_pv8_est$SE[11],
                        model2_pv9_est$SE[11], model2_pv10_est$SE[11])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[11], model2_pv2_est$coef[11], model2_pv3_est$coef
                          [11], model2_pv4_est$coef[11],
                        model2_pv5_est$coef[11], model2_pv6_est$coef[11], model2_pv7_est$coef
                            [11], model2_pv8_est$coef[11],
                        model2_pv9_est$coef[11], model2_pv10_est$coef[11]))

se_coef12_est <- mean(model2_pv1_est$SE[12], model2_pv2_est$SE[12], model2_pv3_est$SE[12],
    model2_pv4_est$SE[12], model2_pv5_est$SE[12],
                    model2_pv6_est$SE[12], model2_pv7_est$SE[12], model2_pv8_est$SE[12],
                        model2_pv9_est$SE[12], model2_pv10_est$SE[12])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[12], model2_pv2_est$coef[12], model2_pv3_est$coef
                          [12], model2_pv4_est$coef[12],
                        model2_pv5_est$coef[12], model2_pv6_est$coef[12], model2_pv7_est$coef
                            [12], model2_pv8_est$coef[12],
                        model2_pv9_est$coef[12], model2_pv10_est$coef[12]))

se_coef13_est <- mean(model2_pv1_est$SE[13], model2_pv2_est$SE[13], model2_pv3_est$SE[13],
    model2_pv4_est$SE[13], model2_pv5_est$SE[13],
                    model2_pv6_est$SE[13], model2_pv7_est$SE[13], model2_pv8_est$SE[13],
                        model2_pv9_est$SE[13], model2_pv10_est$SE[13])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[13], model2_pv2_est$coef[13], model2_pv3_est$coef
                          [13], model2_pv4_est$coef[13],
                        model2_pv5_est$coef[13], model2_pv6_est$coef[13], model2_pv7_est$coef
                            [13], model2_pv8_est$coef[13],
                        model2_pv9_est$coef[13], model2_pv10_est$coef[13]))

se_coef14_est <- mean(model2_pv1_est$SE[14], model2_pv2_est$SE[14], model2_pv3_est$SE[14],
    model2_pv4_est$SE[14], model2_pv5_est$SE[14],
                    model2_pv6_est$SE[14], model2_pv7_est$SE[14], model2_pv8_est$SE[14],
                        model2_pv9_est$SE[14], model2_pv10_est$SE[14])+(1+(1/M))*var(
                      c(model2_pv1_est$coef[14], model2_pv2_est$coef[14], model2_pv3_est$coef
                          [14], model2_pv4_est$coef[14],
                        model2_pv5_est$coef[14], model2_pv6_est$coef[14], model2_pv7_est$coef
                            [14], model2_pv8_est$coef[14],
                        model2_pv9_est$coef[14], model2_pv10_est$coef[14]))


model2_se_est <- rbind(Intercept=se_coef1_est, MALE=se_coef2_est, UNDREM=se_coef3_est,
    METASUM=se_coef4_est, JOYREAD=se_coef5_est,
```

```
                    RESILIENCE=se_coef6_est, SCREADCOMP=se_coef7_est, SCREADDIFF=
                        se_coef8_est, GFOFAIL=se_coef9_est,
                    WORKMAST=se_coef10_est, MASTGOAL=se_coef11_est, WHICH_AREA=
                        se_coef12_est, PUBLIC=se_coef13_est, SCHOOL_ESCS=se_coef14_est)


## t-values
model2_t_est <- model2_coef_est/model2_se_est

## p-values
model2_p_est <- round(2 * pnorm( abs(model2_t_est), lower.tail=FALSE), 3)


## ICC
model2_ICC_est <- mean(model2_pv1_est$ICC, model2_pv2_est$ICC, model2_pv3_est$ICC,
    model2_pv4_est$ICC, model2_pv5_est$ICC,
                    model2_pv6_est$ICC, model2_pv7_est$ICC, model2_pv8_est$ICC,
                        model2_pv9_est$ICC, model2_pv10_est$ICC)

# School variance
mean(model2_pv1_est$vars[1], model2_pv2_est$vars[1], model2_pv3_est$vars[1],
    model2_pv4_est$vars[1], model2_pv5_est$vars[1],
     model2_pv6_est$vars[1], model2_pv7_est$vars[1], model2_pv8_est$vars[1],
         model2_pv9_est$vars[1], model2_pv10_est$vars[1])
# Student level variance
mean(model2_pv1_est$vars[2], model2_pv2_est$vars[2], model2_pv3_est$vars[2],
    model2_pv4_est$vars[2], model2_pv5_est$vars[2],
     model2_pv6_est$vars[2], model2_pv7_est$vars[2], model2_pv8_est$vars[2],
         model2_pv9_est$vars[2], model2_pv10_est$vars[2])



## AIC
model2_loglik_est <- mean(model2_pv1_est$lnl, model2_pv2_est$lnl, model2_pv3_est$lnl,
    model2_pv4_est$lnl, model2_pv5_est$lnl,
                       model2_pv6_est$lnl, model2_pv7_est$lnl, model2_pv8_est$lnl,
                           model2_pv9_est$lnl, model2_pv10_est$lnl)

model2_AIC_est <- -2 * model2_loglik_est + 2 * 14

## BIC
model2_BIC_est <- log(nrow(pisa2018final_EST)) * 14 - 2 * model2_loglik_est
```