Oaxaca Blinder Decomposition of Gender Differences in Reading: An Analysis of Finland and Norway in PISA 2018

Andrei-Iulian Chitic

Master of Science (Assessment, Measurement and Evaluation)

30 credits

Centre for Educational Measurement

Faculty of Educational Sciences

Spring 2022

Contents

List of Tables

List of Figures

Acknowledgment

I want to thank each member of the Centre for Educational Measurement. I am forever grateful that I had the opportunity to pursue this master's degree. To my supervisors, Rolf Vegar Olsen and Henrik Galligani Ræder, I am entirely indebted. Your encouragement, support, and expertise have meant a lot. I consider myself privileged that I had the chance to work with you. I am grateful to my former teachers. All of you have contributed to my success. I will never forget that.

Many thanks to my kind classmates. I have learned a lot of things from you all. Because of you, I have become a better version of myself, and I consider myself blessed to be your classmate.

Hats off to Maria for the helpful comments and proofreading (and the encouragement pictures!).

Pentru cea mai deosebită persoană... mama, tu meriți cel mai special mulțumesc. Lucrarea această de disertație este dovada că mi-am ținut promisiunea față de tine făcută cu un an în urmă în aeroport. Mulțumesc că ai fost mereu lângă mine și m-ai susținut continuu, fără susținerea ta nu aș fi putut absolvi. Tu m-ai învățat mereu să țintesc cât mai sus. Nu am cuvinte să descriu cât însemni pentru mine..

Sper că te-am făcut mândră.

Această lucrare va rămâne permanent în arhiva universității din Oslo ca martoră a muncii noastre asidue.

Popular Abstract

Disparities in reading between girls and boys are ordinary in International Large-Scale Assessments (ILSA). ILSAs have shown that males perform worse than their female counterparts in reading. Program for International Student Assessment (PISA) is one of the ILSAs that focuses on assessing the performance of 15 years old students in 3 distinct domains reading, science, and mathematics. In this study, data from PISA 2018 reading section are used to examine the gender differences in reading achievements taking into consideration the following variables: reading enjoyment, teachers' stimulation of reading engagement, motivation to master tasks, self-concept in reading, perceived difficulty in reading, frequency of reading enjoyment. Applying the so-called Oaxaca-Blinder decomposition (OBD), this study found that the reading enjoyment predictor accounted for most of the explained gap in reading in both countries.

Abstract

The complex design of International-Large Scale Assessments (ILSAs), such as the Program for International Student Assessment (PISA), allowed for examining the gender differences. Using the latest PISA 2018 reading literacy data and the theoretical framework of self-determination theory (SDT) suggested by Deci and Ryan (1985), this study examined the gender differences in reading in Norway and Finland. This study answered the main question of how boys would perform if they possessed similar characteristics as girls. With the help of the Oaxaca-Blinder decomposition (OBD) technique, this question was answered. Using predictors of reading behavior such as reading enjoyment, teachers' stimulation of reading engagement, motivation to master tasks, the positive-self concept in reading, perceived difficulty in reading, and frequency of reading enjoyment. It was found that if boys had possessed similar reading enjoyment levels as girls, the gap in both countries would be reduced. For Norway, the gap would be reduced by 4.4 points out of 46.8 points, and in Finland, the reduction would be 11.9 points out of 48.9. This study offered educational stakeholders information to reduce the gap between genders in reading achievement.

Keywords: PISA, Oaxaca-Blinder decomposition, reading achievement, Norway, Finland

Chapter 1 Introduction

1.1 The Importance of Reading

Reading is necessary to perform well in any situation because meaning, in most cases, is communicated through written texts. Reading is a complex activity. In addition, to decoding the letters and words and understanding the content, the reader also needs focus, effort, perseverance, and determination to initiate, maintain and complete the activity (Olastuen, 2018). Numerous International Large Scale Assessments such as PIRLS (Progress In International Reading Literacy Study), PISA (Programme for International Student Assessment), and EGRA (Early Grade for Reading Assessment) have an interest in assessing this skill.

1.2 Gender Differences in Reading

Across the sphere of education, 21st-century society is driven to address equity issues between the sexes (UNESCO, 2011). According to The Nation's Report Card (2019), significant and large gender differences are consistently found in ILSAs. Gender differences in reading are highly usual, showing that girls outperform boys (Schulz-Heidorf & Støle, 2018). Gender differences frequently occur in adolescent samples rather than in younger samples (Torppa et al., 2018). Studying gender differences in reading over time has shown that gender differences arise during middle school and shrink until grade 12 (Schulz-Heidorf & Støle, 2018). PISA's data offer an opportunity to analyze gender differences; from the first assessment in 2000, Organisation for Economic Co-operation and Development (OECD) reported substantial gender differences in many countries (Zehner et al., 2018).

Reading is an important activity that represents the main factor for educational achievement; discussing boys' underachievement has become the central theme of assessment

and academic research. Some studies even highlighted that boys are considered reading impaired (Mullis et al., 2012; Torppa et al., 2018). The main reason for gender differences in reading achievement is the gap between boys' and girls' reading behaviors and how much time they devote to reading for enjoyment (Scholes, 2019). In general, in PISA, girls performed better across all subscales in the reading with more significant differences reported for more complex skills; children reading behaviors and attitudes are thought to play a role in this situation (Rosén et al., 2022).

An extensive range of studies has found that in more gender-equal societies, girls tend to have a higher reading advantage (Guiso et al., 2008; Reilly, 2012). Based on data from 42 countries in PISA, van Langen et al. (2006) discovered that girls' advantage in reading is prominent in education systems with low levels of differentiation. According to a previous OECD study (2016), there would be a 23-point difference in the reading achievement scores for the average OECD country if boys enjoyed reading as much as girls did. A recent systematic review concluded a higher percentage of frequent, avid, and motivated readers among girls than among boys (Jabbar & Warraich, 2022).

1.3 Reading Literacy in PISA 2018 Framework

Throughout the PISA development, the definition of reading literacy over the years has not been substantially changed. For the 2018 assessment, reading literacy was described as:

”...An individual’s capacity to understand, use, evaluate, reflect on, and engage with texts to achieve one’s goals, develop one’s knowledge and potential, and participate in society.” (OECD, 2019d, p. 28). This definition resulted from pivotal developments in the domain specification and considering the evidence of numerous empirical studies. The assessment used the subscales utilized in 2000 and 2009, being renamed "locating information," "understanding," and "evaluating and reflecting"(OECD, 2019). The PISA 2018 reading

literacy framework incorporates components of the 2009 and 2015 frameworks that are still applicable to PISA 2018; reading literacy is viewed as goal-driven, analytical and intertextual within PISA 2018 (OECD, 2019d).

The reading literacy assessment aimed to scan and analyze the level of proficiency of 15-year-old students since they are about to complete their mandatory education (OECD, 2019d). PISA's reading skills concept requires a succession of circumstances where students read and differentiate how specific texts such as news, memos, printed books, and online boards were presented. Additionally, it entails readers to use different procedures when they approach a variety of limited to comprehensive texts.

Table 1

Illustration of the 2018 Process Typology to 2018 Reporting Scales

2018 Cognitive processes	Superordinate Category Used for Scaling in 2018
Reading fluently	Reported on the PISA scale
Accessing and retrieving information within a text	Locating information
Searching for and selecting relevant text	
Representing literal meaning	Understanding
Integrating and generating inferences	
Assessing quality and credibility	Evaluating and Reflecting
Reflecting on content and form	
Detecting and handling conflict	

Note. Reprinted from PISA 2018 Reading Framework. 21–71 (p. 37) by OECD (2019d).

The authors split the texts into two categories, single texts and multiple texts, each of the texts has distinct tasks (OECD, 2019e). The PISA 2018 framework is centered on three

reading competencies: Locating information, understanding and evaluating, and reflecting, with an even more fine-grained structure presented in **Table 1**. The classification of the texts is acquired from Werlich (1976)**,** which differentiates among description texts, narration texts, exposition texts, argumentation texts, instruction texts, and transaction texts; each of the texts has the purpose of assessing students' reading literacy.

1.4 OECD and The Nordic Model

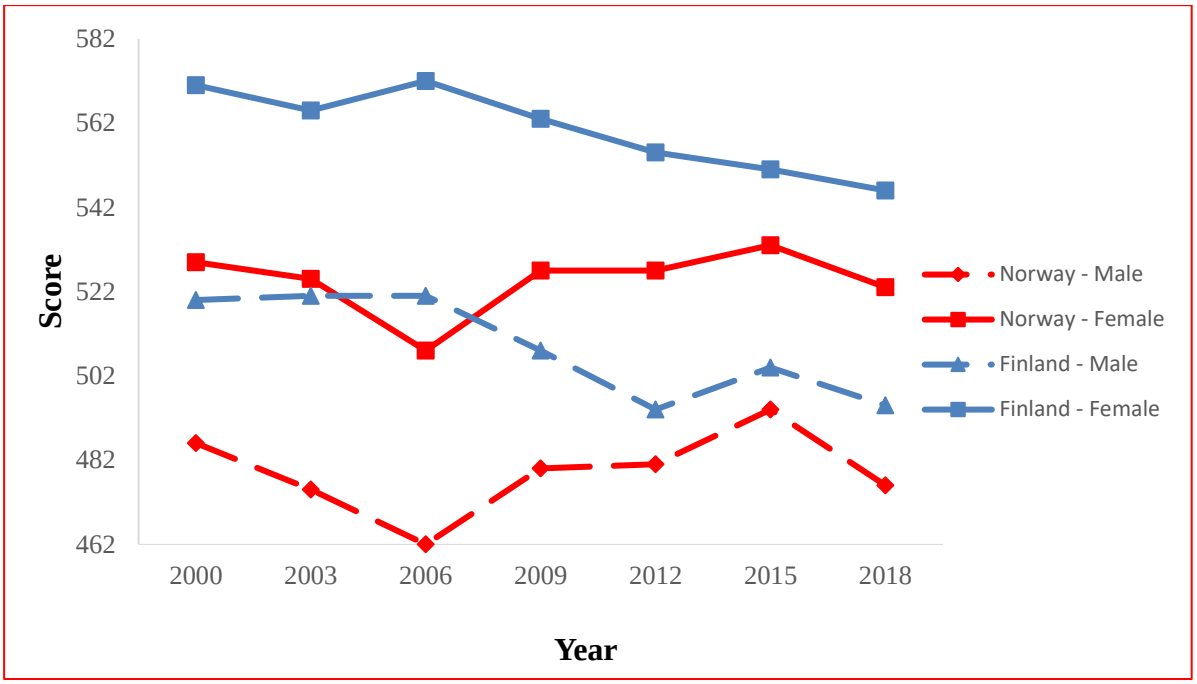
The Nordic Model is a concept often used to encapsulate a set of specific features of Nordic societies, such as: "a long democratic tradition, a regulated economic regime, a socially founded redistribution of benefits', a strong and relatively undivided union movement, a solid social-democratic party" (Larsen et al., 2021, p. 6). Therefore, Nordic countries are seen as countries with high educational opportunities, equality, and equity (Frønes et al., 2020). Nordic countries have been described as progressive and pupil-centered, underlining pupil activity in the tradition of John Dewey and the German concept of arbeitsschule (Larsen et al., 2021).

Despite all those findings, ILSAs cannot completely confirm that the Nordic educational systems live up to the ideals of the Nordic model. Instead, the observed large gap in reading between boys and girls appears as a contradiction to the Nordic model. Although there are differences in how education is governed in Norway and Finland, they share the same fundamental characteristics with free public comprehensive the schools until students are approximately 15-years old.

1.4.1 Gender Trends in PISA Reading in Norway and and

Figure 1

Average Reading Achievement by Gender for Finland and Norway in PISA 2000-2018.



Note. The gender difference is statistically significant for all years.

OECD (2022) historical data was used to create **Figure 1**. Finland's and Norway's gender gap in reading has consistently been the highest across all countries participating in PISA. However, it has been relatively stable. Given the ambitions expressed in the Nordic model, this finding can be perceived as surprising. To be more specific, girls have consistently performed better than boys in reading throughout Norway's and Finland's history in PISA. Finland managed in 2009 to have the highest gender gap in reading out of all OECD countries (55 points). In addition, this issue repeated in 2018 when the gap between Finnish girls and boys reached 52 points, the highest out of all OECD countries (OECD, 2019a, 2019b). Over the period that PISA has existed, the gender gap in Norway has also been very high and among the largest in the OECD area.

1.5 Self-Determination Theory

According to self-determination theory (SDT), students are viewed as living organisms that act on the environment rather than passively recipients of the information. Motivation and personality are studied under SDT (Deci & Ryan, 1985). Based on the SDT, students

are said to have three universal psychological needs: competence, autonomy, and relatedness (Deci & Ryan, 1985). Their level of independence distinguishes different extrinsic motivations. According to Ryan and Deci (2000b), amotivated students feel they lack the ability to achieve a goal while dealing with the activity; they do not assign any value to it. Their level of competence is low; the external regulation is considered to be among minor autonomous types of motivation (Jeno et al., 2017). In the case of externally regulated students, they perform an activity to receive a reward or avoid punishment (Deci & Ryan, 2004). Students' social context inhibits their psychological needs or supports them; thus, the SDT highlights that supportive interpersonal contexts are critical (Jeno et al., 2017).

As Reeve (2009) affirmed, when teachers consider student perspectives and provide choices and opportunities for students and nurture their motivational resources, students' needs for autonomy, competence, and relatedness are likely to be met; as a result, self-determination is promoted. Students who have teachers who support their autonomy have greater self-esteem and confidently believe they are more competent than others (Deci et al., 1981; Diseth et al., 2012). Motivating oneself intrinsically, through self-determination, is the most optimal form of motivation; it describes the act of investing oneself in an activity based on one's interests, pleasure, and overall satisfaction (De Naeghel et al., 2012). It reflects "the inherent tendency to seek novelty and challenges" (Ryan & Deci, 2000a, p. 70).

When it comes to reading, children pursue the reading activity because they find it enjoyable; however, although children are not spontaneously interested in reading, a degree of autonomy or self-determination is still possible (De Naeghel et al., 2012). If children’s self-concept is not supported by an external factor (e.g., parents, teachers), they will not view themselves as dedicated learners (Brumfit, 1993). This theory supports the strong relationship between self-concept and reading achievement (De Naeghel et al., 2012). Hence, this thesis

explores the gender differences in reading through an amalgam of concepts closely related to self-determination theory.

1.5.1 Reading Behaviors Viewed as Influential Factors

Reading behaviors are very relevant potential predictors. Whether positive or negative, reading behaviors have a self-reinforcing effect on reading comprehension. Those who read a lot and have positive attitudes become better readers, while those who read less will progress slower (Olastuen, 2018). Engagement can be both about the desire to read for the sake of pleasure and about the driving force that arises when one sees the importance of understanding a text; commitment is also necessary to keep concentration up when the text is demanding, and the reader must read accurately and concentrated on getting the content (Roe & Jensen, 2017).

Hence, it is confirmed that those factors can influence reading achievement. Due to the complex design of PISA, this study managed to incorporate only the following reading predictors: reading enjoyment, teachers' stimulation of reading engagement, motivation to master tasks, positive-self concept in reading, perceived difficulty in reading, and frequency of reading enjoyment. The upcoming subchapters describe in detail the predictors mentioned above. However, this is not an exhaustive list of possible and essential predictors. There could be a possibility that those factors are not the only ones affecting reading achievement, such as self-efficacy, reading strategy, perseverance in reading activities, various forms of reading motivation attributes, etc.

1.5.2 Reading Enjoyment

Prior research indicates that reading enjoyment correlates positively with reading achievement in countries with high academic achievement (Malanchini et al., 2017). Cheema (2018) finds a positive association between reading enjoyment and academic performance in high-performing countries and a negative relation in low-performing countries. Reading

enjoyment has been linked to comprehension and cognition, is a measure of reading involvement and interest, and is part of intrinsic motivation that can influence achievement in reading (Guthrie et al., 2007). Comparatively, girls are more likely to enjoy reading (Cheema, 2018; Loveless, 2015). In one of the OECD's studies, girls in all countries read for enjoyment more than the boys; one exception was Korea (OECD, 2010).

Another study suggested that reading enjoyment is also a prerequisite for successful reading comprehension (Zasacka & Bulkowski, 2017). Therefore, one must keep in mind that reading enjoyment was linked not only to reading comprehension but also to students' overall school performance (Rojas-Torres et al., 2021). A positive relationship has been found between perceived reading ability and enjoyment of reading (Clark & Rumbold, 2006). Reynolds and colleagues (2009) found that knowledge test performance relates to reading enjoyment and perceived competence. Shiel and Cosgrove (2002) found that the frequency of reading for enjoyment was one of the factors which explained reading performance.

1.5.3 Teachers' Stimulation of Reading Engagement

Students must be instilled with a love of reading (Cambria & Guthrie, 2010). To encourage adolescents to read for pleasure and self-determination, several methods can be used: stimulating their intrinsic motivation, developing their confidence, and allowing them to see the value of reading; teachers play a crucial role in developing students' reading behaviors (Ruddell, 1995). Research has shown that teachers' emotional support is related to increased engagement in learning, academic enjoyment, and self-efficacy, leading to better student effort and persistence (Federici & Skaalvik, 2013; Ruzek et al., 2016; Sakiz et al., 2012). In addition, higher intrinsic motivation and lower anxiety levels were associated with teachers' support (Erdol & Arikan, 2022; Yu & Singh, 2018). Wray and colleagues (2000) have discovered that teachers' appreciation of their reading and writing skills directly impacts students' performance.

Feedback on each student's performance can enable them to achieve better outcomes (Hattie & Timperley, 2007; Tunstall & Gsipps, 1996). According to the OECD, students worldwide whose teachers support most of their language lessons have significantly higher scores than those who do not (Munita, 2016; OECD, 2019c). We expect a similar finding as in Rojas-Torres et al. (2021); more specifically, teachers' stimulation of reading engagement predictor is not highly related to students' performance since many factors have influenced the students throughout their lives. Hence, a modest positive relationship between teachers' stimulation of reading engagement and reading achievement is expected.

1.5.4 Motivation to Master Tasks

According to Schunk et al. (2012), motivation can be crucial to how well students work on given tasks; motivation is essential to get something done or learn something. Students with high motivation will start with lessons and activities different from students with low motivation, and increased motivation is the best basis for learning (Zimmerman, 2000). According to Marshall (1987), students' motivation is beneficial to the learner. Motivated students seek help more often if there is something they do not comprehend, take notes, organize and practice what to learn more often than students with lower motivation levels (Schunk et al., 2012). Bandura (1997) introduced the concept of self-efficacy. Self-efficacy is an individual's belief in their ability to succeed under a particular circumstance (Bandura, 1997).

Norwegian media barometer, the Student Survey, and Ungdata concluded from interviews and multiple analyses that Norwegian students read a little and are not incredibly motivated to read (Olastuen, 2018). In addition, results from ILSAs such as PISA and PIRLS indicate a paradox; despite having a low level of motivation, the Finnish students are still high achievers in reading (Ng & Bartlett, 2017). As a result of concerns about students' lack of reading motivation, the Ministry of Education and Culture in Finland developed the joy of reading literacy program (Ng & Bartlett, 2017).

1.5.5 Self-Concept in Reading

Self-concept scales developed by OECD measured students' perception of competence and difficulty in reading. Numerous researchers identified self-concept as a multidimensional concept. Combs (1962) describes the self-concept as "A person's idea of himself or herself." Marsh (1986) specified that the association between self-concept and reading has three components, more precisely: perception of difficulty, perception of competence, and the reading attitude; Nicholls and Miller (1983) differentiate between the perceptions above; according to them, the perceptions of difficulty and the perceptions of competence do not have to agree with each other. An extensive range of studies answer the question of "How does self-concept affect achievement in reading?" Petscher's (2010) meta-analysis identified that reading behaviors such as reading self-concept and reading achievement are moderately correlated.

Chapman and Tunmen (1995) and Marsh et al. (1985) have concluded a positive correlation between self-concept in reading and reading achievement. Positivity toward reading is typically associated with higher reading achievement levels; inversely, negativity toward reading is associated with lower reading achievement levels (Rider & Colmar, 2005). This could mean that students who have a favorable opinion about their reading ability perform better, whereas students who think negatively about their reading ability perform worse. Finally, Mullis and colleagues (2012) have found that reading motivation will generally be higher in students who believe they can succeed in reading and therefore expect to succeed. According to Rider and Colmar (2005), there is evidence that girls tend to have a stronger reading self-concept than boys and display more positive reading behaviors toward reading.

1.6 The Present Study

This study examines how measures of reading behaviors collectively explain the gender gap in reading in PISA 2018 in Norway and Finland. In a previous study, Sari (2020) applied a

similar approach using Oaxaca-Blinder decomposition to decompose the reading gap in PIRLS. The present study complements the abovementioned study by applying Oaxaca-Blinder decomposition to a different age group. Hence, the data, the age of participants, and the variables are distinct. It is worth mentioning that the present study focuses only on Norway and Finland compared to the previous research, which targeted all the Scandinavian countries.

Likewise, for its predecessor, various statistical techniques are employed in this study to illustrate how dissimilar reading behaviors predictors concerning reading achievement add to the gender gap in achievement. This topic is timely, relevant, and essential, primarily since it refers to possible educational inequalities. Shedding light on the specifics of the gender gaps in reading, especially in light of other reading-related constructs, can generate relevant knowledge to understand these gaps.

This study aims to enrich the existing knowledge about gender differences. This study seeks to measure the gender gap in reading achievement, considering students' characteristics. Then this study utilizes Oaxaca-Blinder decomposition to decompose the gap identified in reading into two distinct parts. Students' factors can explain one part, and one part is unexplained, and it results from differences in the predictor variables' associations with the outcomes in the two groups (Sen, 2014). It is expected that this study will provide insights into gender differences in reading in PISA.

This study offers educational stakeholders information to reduce the gap between genders in reading achievement. Finally, this study examines potential sources of gender differences. In the Methods chapter, we will go over the techniques, procedures, and variables in further detail. Using the results from PISA 2018 for Norway and Finland, the assessment is done taking into consideration the importance of the following predictors: reading enjoyment,

teachers’ stimulation of reading engagement, motivation to master tasks, the positive-self concept in reading, perceived difficulty in reading, frequency of reading enjoyment.

This study focuses on answering those research questions :

RQ1) To what extent are the variables that measure students' attributes associated? (reading enjoyment, teachers’ stimulation of reading engagement, motivation to master tasks, self-concept, perceived difficulty in reading, frequency of reading enjoyment)?

RQ2) To what extent are the students' attributes linked with their reading achievement?

RQ3) To what extent are girls and boys different concerning reading behaviors?

RQ4) How large is the average gender gap after controlling for reading behavior predictors?

RQ5) How much would boys' reading achievement increase if they had similar reading behaviors as girls?

A specific characteristic of the method used in this study is counterfactual thinking. Counterfactual thinking emerges early in life and seems apparent as soon as a person is proficient; counterfactual thinking is illustrated using subjunctives such as "if only" (Neimeyer et al., 2021). A common counterfactual concept is imagining how the past might have changed (Roese et al., 2017). By imagining how antecedent circumstances or consequences could have been different, one can undo a current event that has already taken place (Maitner & Summerville, 2022).

Elder et al. (2009) suggested that the counterfactual scenario could be facilitated by Oaxaca-Blinder decomposition by estimating separate linear regressions for the two groups when the gap in the mean outcomes between the groups mentioned above is present. Hence

Oaxaca-Blinder decomposition allows us to distinguish between explained and unexplained components of the gap.

The study is organized as follows: Chapter 2 describes the analysis methodology and techniques. Chapter 3 contains the results, and lastly, Chapter 4 comprises the discussions, limitations, further research, and conclusions.

Chapter 2 Methods

First, the data and sample of the students are described, followed by a detailed description of the variables included in the analysis. Next, the subsequent procedures used for the study will be described in detail (Oaxaca-Blinder decomposition and multiple regression). Lastly, the data missingness is to be discussed.

2.1 Data and Sample

The paper analyses data from the PISA 2018 main study. All participating students were 15 years old when the data was collected. The total number of students who participated in PISA 2018 in Norway and Finland altogether was 11462[1]. In both countries, girls constituted 51% of the respondents in this study, while boys comprised 49%. Separately each sample consisted of under 3000 students per country per gender.

It is common for PISA to use a two-stage sampling procedure, where first a set of schools are sampled, and then a sample of 30-35 students is drawn within the sampled schools. PISA 2018 data is available online and can be downloaded from OECD's website (https://www.oecd.org/pisa/data/2018database). The prospective analysis is done using participants' total weight and variables such as reading enjoyment (*joyread*), teachers'

1. Appendix C, **Table C1** provides a detailed overview of the samples.

stimulation of reading engagement *(stimread*), motivation to master tasks (*workmast*), self-concept in reading (*scir*), perceived difficulty in reading (*pdir*), frequency of reading enjoyment *(readjoyfreq*).

2.2 Sampling Design, Weight, and Plausible Values in PISA

Stratified multistage sampling is used in PISA. Schools are selected by probability proportional to their school enrollment, which would lead to an approximately equal probability of all students being selected. However, all countries use stratified sampling to ensure good coverage of the specific subgroups of students (e.g., geographical location, minority); this results in a non-equal probability of being selected. To account for that, PISA uses weights (OECD, 2019a). Upon sampling, students are assigned a weight that tells how much each student is reflected in the population; school weight and student-within-school weight are incorporated together when calculating the final student weight (W_FSTUWT) (OECD, 2019a).

Moreover, In PISA, sample variance is addressed by a method named balanced repeated replication (BRR). This method is similar to the jackknife method used by other studies like PIRLS and TIMSS. (Rust & Rao, 1996; Ward, 2020). According to Wu (2005), Students' unobservable latent achievements are represented by plausible values. Crato (2021) noted that:

> "...To increase accuracy, these ILSA use plausible values (multiple imputations) drawn from a posteriori distribution which is constructed by combining the IRT scaling of the test items with a latent regression model with information from the student context questionnaire within a population model; for each student 10 plausible values are computed in PISA since 2015" (p. 251).

The computation of plausible values proves helpful when the magnitude of measurement error is substantial. The nature of ILSA's survey indicates that the focus has to be on the population parameters rather than the individual respondent (von Davier et al., 2009).

2.3 Scaling and Variables

This study uses the variables related to students' behavior from the primary student survey. Using IRT, PISA, Trends in International Mathematics and Science Study (TIMSS), and PIRLS, students' test scores were computed; for the achievement scores, the mean is being standardized approximately to 500, and the standard deviation is approximately 100 (Crato, 2021). The measures based on students' self-report were standardized across Norwegian and Finnish samples, whereas the measures developed by OECD were standardized across all OECD countries.

2.4 Development of Scales

This study uses different variables to understand the gender differences in reading. Most of the variables were continuous, except gender. The scales are standardized for the sample used in this study, where the mean and standard deviation is defined to be 0 and 1, respectively.

2.4.1 Scales Developed by OECD

Table 2

Summary of Variables and Scales

Concept	Scale	Items
Reading Enjoyment	joyread	5
Teachers' Stimulation of Reading Engagement	stimread	4
Motivation to Master Tasks	workmast	4
Self-Concept in Reading	scir	3
Perceived Difficulty in Reading	pdir	3
Frequency of Reading Enjoyment	readjoyfreq	1

The first scale of this study is *joyread,* which measures reading enjoyment. It was computed using five self-reported items, "I read only if I have to," "Reading is one of my favourite hobbies," "I like talking about books with people," and "For me, reading is a waste of time," "I read only to get information that I need." Ranging from "Strongly disagree" to "Strongly agree," each response category is represented on a four-point Likert scale. Item Response Theory (IRT) scaling was implemented. Higher values on the scale reflect that students read more for enjoyment (OECD, 2018).

The second scale is *stimread*, which measures teachers' stimulation of reading engagement. It was computed based on four items, e.g., "The teacher encourages students to express their opinion about a text," "The teacher helps students relate the stories they read to their lives," and "The teacher shows students how the information in texts builds on what they already know," and "The teacher poses questions that motivate students to participate actively." The statements refer to how often specific activities happened in the language instruction class. The students could choose from "never or hardly ever" to "in all lessons." Higher values on the scale reflect that students perceived their teachers providing more reading stimulation (OECD, 2018).

The third scale *workmast* was constructed out of those three statements, e.g., "I find satisfaction in working as hard as I can"; "Once I start a task, I persist until it is finished"; "Part of the enjoyment I get from doing things is when I improve on my past performance." The responses were coded as a four-point Likert agreement scale: "strongly disagree," "disagree," "agree," and "strongly agree." Students with higher values are more motivated (OECD, 2018).

In our study, *workmast* and *stimread* are not explicitly related to reading. Instead, *workmast* represents a specific aspect of motivation, more specifically, mastery. Whereas *stimread* reflects a measure partly related to reading, it refers to the teachers' actions on how often they stimulate students in the language instruction lessons.

2.4.2 New Scales Developed for this Study

In addition to the scales already developed by the OECD, three more scales were developed based on the student background questionnaire data. The questionnaire consisted of six questions to measure students' self-concept in reading and one question on frequency of reading enjoyment. However, the OECD did not provide scales based on these items. One reason for this might be that some of the items were negatively worded, leading to threats to the construct validity and reliability of the instrument (Chyung et al., 2018). To avoid those issues, the negatively worded items were reversed. Confirmatory Factor Analysis (CFA) was used to ensure the internal validity of the new scales.

Details of the CFA are presented in Appendix C in **Table C3** and **Table C4**. In summary, according to Hu and Bentler (1999) there is decent fit (RMSEA 0.08 > 0.06, SRMR 0.03 < 0.09, CFI 0.955 > 0.95) of the two-factor model. With the possible exception of RMSEA. The factor loadings are high for all items in the two-factor model. Factor loadings for the three items on the difficulty scale (*pdir*) were low in the single factor solution. Analysis of variance (ANOVA) confirms that the two-factor model is significantly better than the single-factor model (Appendix C**, Table C5**). Lastly, inter-item correlations were checked to ensure the validity of the measures based on the self-concept items. The items within the same factor are highly correlated. (Appendix C, **Table C8**).

On the fourth scale *scir*, the students were asked how much they agree or disagree with the three statements about themselves, e.g., "I am a good reader," "I am able to understand difficult texts," "I read fluently,"; The responses were coded as a four-point Likert agreement scale: "strongly disagree," "disagree," "agree," "strongly agree." Higher values on the scale reflect that students possess a higher level of self-concept in reading.

On the fifth scale *pdir*, the students were asked how much they agree or disagree with those three statements about themselves, e.g., "I have always had difficulty with reading," "I

have to read a text several times before completely understanding it," "I find it difficult to answer questions about a text"; The responses were coded as a four-point Likert agreement scale: "strongly disagree," "disagree," "agree," "strongly agree." This time respondents' responses were reversed. The scale indicates the perceived difficulty in reading. More specifically, higher values on the scale mean that the students think of themselves as having fewer reading problems.

On the sixth scale, *readjoyfreq,* the students were asked how much time they usually spend reading for enjoyment. The responses were coded as a five-point frequency scale as follows: "I do not read for enjoyment," "30 minutes or less a day", "More than 30 minutes to less than 60 minutes a day", "1 to 2 hours a day", "More than 2 hours a day[2]"; Higher positive values on the scale mean that students spend a higher amount of time reading for enjoyment than the average student.

2.5 Missing Data

For most measures, the proportion of missing values is relatively low. Across measures used, the missing response proportion ranges from 10% to 16 (Appendix C, **Table C2**). Before proceeding with the analysis, the pairwise deletion method was used to handle the data missingness. Compared to the listwise deletion, the main advantage of the pairwise method is preserving more data (Kim & Curry, 1977).

2.6 Procedures

The analyses were conducted using The International Association for the Evaluation of Educational Achievement (IEA) International Database Analyzer (Mikheeva & Meyer, 2019). When carrying out the analysis, the IDB Analyzer produces an SPSS syntax. IDB Analyzer is a free software program developed by the IEA which allows one to analyze and combine the

2. More details can be found in Appendix B (B.1 "Computation of the New Indices "SCIR," "READJOYFREQ" "PDIR").

data from the majority of IEA's studies. The software appropriately uses all ten plausible values for the reading score. Proper weights and procedures are employed to estimate robust sampling variance, resulting in proper estimates of standard errors being reported.

Firstly, a mean analysis by gender was performed on the selected variables and achievements. After that, Pearson's correlation was calculated. Additionally, Cohen's d was estimated to determine the effect size of the mean differences, while t-tests were used to detect if significant differences exist between groups.

Up next, multiple linear regression was employed. Oaxaca-Blinder decomposition was computed using the regression coefficients and the variables' mean scores acquired from prior procedures. The upcoming sections focus on estimating the regression, followed by the Oaxaca-Blinder decomposition technique.

2.7 Regression Estimation

This study examines how students' background characteristics (reading behaviors) collectively predict reading proficiency. It is assumed that the linearity assumption is met between the explanatory variables and the outcome. The following equation is adapted from (Bangdiwala, 2018).

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_j X_{ji} + \varepsilon_i \quad (1)$$

Y_i represents the dependent variable (reading achievement) of the student "i."

β_0 represents the intercept (constant)

β_j represents the coefficient of the regression of variable "j."

X_{ji} represents the student's "i" variable of "j."

ε_i represents the error (residual) of student "i."

i =index of students (1 to n)

j =index of variables (1 to k)

The multilinear regression analysis over the combined sample of male and female students from Norway and Finland used gender as a categorical variable for the first step. For the forthcoming decomposition, regression models are conducted separately for the girls and boys to check for differences. Lastly, R^2 will be outlined, and to test for multicollinearity Variance Inflation Factor known as (VIF) is used.

2.8 What is Oaxaca-Blinder Decomposition?

Oaxaca-Blinder decomposition was initially used to study the earning differentials across groups, e.g., gender, socioeconomic background, ethnicity, and public versus private. This gave a tool to trial the efficacy of various work-related programs (Barrera-Osorio et al., 2011). Oaxaca-Blinder decomposition is a regression-based approach that divides the gap into two parts. First, the explained portion of the gap is generally referred to as endowments (the part reflecting the group differences in mean scores on the predictors). The unexplained part is associated with differences in the regression coefficients, and we will generally refer to those as the coefficient terms.

After the disparities in the reading achievement have been identified, the upcoming step consists of determining the possible reasons behind the differences. For example, the Oaxaca-Blinder decomposition seeks to respond to questions like, what if boys were more like girls? What if boys possessed similar characteristics as girls? How much better could they be predicted to read?

The gender disparities will be dissected using the decomposition, assessing their contributing factors. It is intended to explain the discrepancies in the reading achievement (outcome) between the two groups through the explanatory variables (contributing factors). Using the observed contributing factors, the difference in the mean of the two groups is computed, resulting in the actual counterfactual components. The objective of Oaxaca-Blinder

decomposition in this study is to elucidate the variables (contributing factors) that are more relevant in explaining the gender differences in reading achievement and decompose the gap by gender.

As Jann (2008) stated, the decomposition uses the regression analysis model, fulfilling the assumptions of linearity and the conditional mean equals zero. (Fortin et al., 2011). According to Hlavac (2014), the OBD technique is adequate for exploring the disparities in any outcome variable that exhibits cross-group differences. There is a substantial body of literature on gender effects and reading achievements, but limited research uses the decomposition method to break down the apparent gender difference in reading. Multiple studies have been carried out by using PISA datasets to study differences between other groups or between countries (Ammermueller, 2007; Barrera-Osorio et al., 2011; Botezat Alina, 2012; Gevrek & Seiberlich, 2014; Giménez & Aristizábal, 2017; Nieto & Ramos, 2015; Ning et al., 2016).

With the help of multilinear regression, every group's mean outcome could be decomposed by multiplying the mean of the contributing factors by the regression's coefficient. The total disparity can be attributed to either a) disparities in the X's, the explained component, or b) disparities in the unexplained component (β's). Lastly, the final OBD formula is adapted from (Jann, 2008).

$$\bar{Y}_G - \bar{Y}_B = \underbrace{(\bar{X}_G - \bar{X}_B)'\hat{\beta}_B}_{\text{Endowments}} + \underbrace{\bar{X}_B'(\hat{\beta}_G - \hat{\beta}_B)}_{\text{Coefficients}} + \underbrace{(\bar{X}_G - \bar{X}_B)'(\hat{\beta}_G - \hat{\beta}_B)}_{\text{Interaction}} \quad (2)$$

In the above formula, the estimates of $\hat{\beta}_N$ comes from the multilinear regression conducted for the two groups separately. Next, the mean of the contributing factors (predictors) samples is combined to compute the OBD decomposition. The raw disparities in reading scores between two groups (males and females) were decomposed using this

method, resulting in unexplained and explained parts of the discrepancy. This method, as Jann (2008) stated, is known as the "threefold Oaxaca-Blinder decomposition of the mean outcome difference" (p. 2).

As illustrated previously, the OBD is computed from the perspective of the male group (Jann, 2008; Kröger & Hartmann, 2021). By keeping the coefficient constant, the endowment component measures the anticipated change in reading achievement of the male group if it achieved predictor levels similar to the female group. Identical to the previous studies, the actual study uses the counterfactual scenario. By looking at how much the reading scores of the male group would change if the attributes of males were equalized to that of the female group, we can ascertain the contribution of their characteristics (our contributing factors) (Olsen & Blömeke, 2018).

Jann (2008) mentioned that the interaction component represents the difference in both coefficients and endowments between the male and female groups, which is being taken into account. Jann (2018) describes it as the part that remained after controlling for coefficients and endowments, which is generally insignificant and small. The discussion of this study is mostly around endowments, as Olsen and Blömeke (2018) stated that direct interpretations of the separate coefficient terms are challenging.

Chapter 3 Results

This chapter commences with an illustration of the correlation analysis results. We report the results of the descriptive statistics more specifically (means, standard deviations, and standard errors) for the reading achievements in Norway and Finland. We then discuss the gender disparities in the contributing factors (predictors). Afterward, using a pooled regression of Norway and Finland samples and dummy-coding gender as an explanatory variable, multiple regression analysis is used to model student reading achievement as a function of the

contributing factors used in this study (having the female group as a group of reference). Lastly, the threefold Oaxaca-Blinder decomposition breaks down the gender gap in the reading achievements.

3.1 Relations Among Variables and Reading Achievement

This subsection commences by illustrating the bivariate correlations in Norway and Finland for the contributing factors (predictors) with the mixed sample (male and female) from the PISA student questionnaire. Next, the correlations between reading achievement and contributing factors (predictors) are illustrated.

Table 3

Correlations between Scales from the Student Questionnaire

Variables	1	2	3	4	5	6
1. joyread	—					
2. stimread	0.16	—				
3. workmast	0.20	0.19	—			
4. scir	0.35	0.14	0.27	—		
5. pdir	0.21	0.07	0.16	0.46	—	
6. readjoyfreq	0.38	0.05	0.09	0.19	0.09	—

Note. The average correlations for Norway and Finland are illustrated. All the correlations shown are significant, $p < 0.01$.

Positive low to moderate correlations were found among all variables in **Table 3**. It was discovered that scales *pdir* and *scir* were modestly associated (r = 0.46), and *joyread* and *readjoyfreq (*r = 0.38), *joyread* and *scir* (r = 0.35), *scir* and *workmast* (r = 0.27);

Table 4

Correlations between Contributing Factors (Predictors) and Reading Achievement

Variables	Reading Achievement	
	Norway	Finland
1. joyread	0.31	0.43
2. stimread	0.13	0.11
3. workmast	0.15	0.28
4. scir	0.41	0.43
5. pdir	0.34	0.35
6. readjoyfreq	0.09	0.20

Note. All correlations were significant, having p values < 0.01.

Table 4 presents the correlations between contributing factors and reading achievement. All the variables positively correlated with reading achievement. However, the differences between those two countries are minor. The correlations are more substantial in Finland. For example, student enjoyment of reading (*joyread*) and students' self-concept in reading have the strongest correlation in Finland (r = 0.43), whereas, for Norway, the correlation is less substantial (r = 0.31).

Students' self-concept in reading (*scir*) in Norway has manifested the highest correlation with the reading achievement (r = 0.40), followed by the students' perceived difficulty in reading (*pdir*), (r = 0.34). The values for the motivation to master tasks (*workmast*) and the frequency of the reading enjoyment (*readjoyfreq*) almost halve in Norway compared to Finland. Lastly, teachers' stimulation of reading engagement (*stimread)* and the frequency of the reading enjoyment of the students (*readjoyfreq)* manifested the weakest correlations in both countries. Those results are helpful for the later analyses.

3.2 Descriptive Results

3.2.1 Gender Differences and Reading Achievements in Norway and Finland

Table 5

Variable Means and the Significance Values across Gender and Countries in PISA 2018

	Norway						Finland					
	Girls		Boys				Girls		Boys			
	Mean (s.e)	SD (s.e)	Mean (s.e)	SD (s.e)	t	E.S	Mean (s.e)	SD (s.e)	Mean (s.e)	SD (s.e)	t	E.S
joyread	-0.22 (0.02)	1.15 (0.02)	-0.80 (0.02)	1.04 (0.02)	**19.6**	0.5	0.10 (0.01)	1.12 (0.01)	-0.60 (0.02)	0.96 (0.01)	**24.9**	0.7
stimread	-0.20 (0.02)	0.95 (0.02)	-0.22 (0.02)	1.03 (0.02)	1.5	0	-0.18 (0.03)	0.91 (0.02)	-0.22 (0.02)	0.98 (0.02)	1.5	0
workmast	0.08 (0.02)	0.97 (0.01)	-0.05 (0.02)	1.06 (0.01)	**4.6**	0.1	-0.22 (0.02)	0.92 (0.01)	-0.41 (0.02)	0.95 (0.02)	**7.3**	0.2
scir	0.13 (0.02)	0.95 (0.04)	-0.01 (0.02)	1.19 (0.05)	**6.3**	0.2	-0.01 (0.02)	0.73 (0.05)	-0.11 (0.02)	1.05 (0.05)	**7.0**	0.2
pdir	-0.09 (0.02)	0.98 (0.01)	-0.10 (0.02)	1.02 (0.02)	0.3	0	0.01 (0.02)	1.00 (0.01)	0.17 (0.02)	0.98 (0.01)	**-5.9**	0.2
readjoyfreq	-0.01 (0.03)	1.00 (0.03)	-0.17 (0.01)	0.76 (0.02)	**6.6**	0.2	0.26 (0.02)	1.25 (0.02)	-0.08 (0.02)	0.86 (0.03)	**11.8**	0.3
Reading	**523** (2.6)	96 (1.5)	**476** (2.6)	109 (1.7)			**546** (2.3)	100 (1.5)	**495** (2.9)	92 (1.6)		

Note. t = t-value from test statistic. E.S = effect sizes. Values in the columns "E.S" were rounded to the nearest .1. Significant mean differences are marked in bold.

Table 5 presents the mean values for boys and girls for each predictor in the regression model and the average reading performance in each country. Furthermore, the table includes the t-values from two-sample t-tests for the differences in the mean values for boys and girls. Finland outperformed Norway's average reading performance within the male and female groups. The highest gap in reading was also found in Finland. Girls' performance was higher by 51 points than their male counterparts. In Norway, the gap between the girls'

performance and boys was 47 points. The gap regarding the standard deviation in Norway is significant, whereas the direction is the opposite in Finland.

Across Norway and Finland, all the differences in mean scores of the predictors were relatively small. Girls scored higher, except for the perceived difficulty in reading (*pdir*), whereas boys scored higher in Finland. In Norway, it has been found that both teachers' stimulation of reading engagement (*stimread*) and perceived difficulty in reading (*pdir)* difference in means between boys and girls were not significant. This finding was similar in Finland. Overall, the effect sizes were relatively small, except for the enjoyment of reading (*joyread*) predictor, which manifested a moderate effect size in both countries.

3.3 Regression Results

Table 6

Multiple Regression Results with Pooled Sample Separated by Gender

	Norway			Finland		
	Pooled	Girls	Boys	Pooled	Girls	Boys
	β (s.e)	β (s.e)	β (s.e)	β (s.e)	β (s.e)	β (s.e)
Intercept	**530.47** (2.85)	**527.53** (2.62)	**485.26** (2.78)	**550.54** (1.93)	**546.22** (2.08)	**510.63** (3.24)
Male	**-27.19** (2.80)			**-30.73** (2.72)		
joyread	**14.36** (1.17*)*	**19.11** (1.66)	**7.64** (1.92)	**19.36** (1.49)	**25.73** (2.01)	**17.09** (2.10)
stimread	**5.50** (1.67)	**6.91** (2.01)	**5.56** (2.17)	-0.25 (1.22)	-0.78 (1.75)	1.81 (1.93)
workmast	2.09 (1.43)	-1.03 (1.94)	**5.65** (1.90)	**11.07** (1.34)	**7.70** (1.92)	**14.39** (2.26)
scir	**26.25** (1.75)	**23.10** (2.65)	**30.15** (2.03)	**19.42** (1.62)	**17.87** (2.60)	**22.73** (2.42)
pdir	**19.57** (1.40)	**19.45** (2.15)	**23.03** (2.30)	**17.08** (1.55)	**15.76** (2.49)	**19.62** (2.09)
readjoyfreq	-4.39 (1.60)	-3.79 (1.84)	-7.65 (2.59)	0.77 (1.23)	-2.45 (1.49)	3.37 (2.38)
R-Square	0.25 (0.01)	0.25 (0.02)	0.21 (0.02)	0.31 (0.01)	0.29 (0.02)	0.25 (0.02)

Note. Significant coefficients are marked in bold.

Table 6 presents the multiple regression analysis, including all the predictors and gender. It has been found that when controlling for all the predictors, girls outperformed boys

with 27 points in Norway and 30 in Finland. In Norway, the coefficients for females have proven to be higher than for males. The coefficients are relatively similar across genders in the two countries. However, some of the coefficients favored the boys, such as motivation to master tasks *(workmast)* and self-concept in reading (*scir*) in Norway. Whereas in Finland, boys have higher values than girls in the majority of the regression coefficients: teachers' stimulation of reading engagement (*stimread*), motivation to master tasks (*workmast)*, self-concept in reading (*scir*), and frequency of reading enjoyment (*readjoyfreq*), perceived difficulty in reading (*pdir)*;

A formal test to check the differences between genders in coefficients has not been carried out. However, given the relative size of the standard errors to the gender differences in coefficients, most of them are likely not significant.

For the pooled regression, Norway's not significant coefficients were motivation to master tasks (*workmast*) and frequency of reading enjoyment (*readjoyfreq*). The strongest in the pooled sample in Finland was the enjoyment of reading (*joyread*). In Finland, the non-significant regression coefficients were teachers' stimulation of reading engagement (*stimread*) and frequency of reading enjoyment (*readjoyfreq*). Regarding the Norwegian pooled sample, the highest value of the regression coefficient was the one of the self-concept in reading (*scir)*.

This finding was similar for the Norwegian female and male samples. In Finland, the female sample's highest regression coefficient of the predictor of reading achievement was reading enjoyment *(joyread*). In contrast, for the Finnish males, the regression coefficient of the self-concept in reading (*scir)* was the highest. In Norway and Finland, the pooled multiple regression models accounted for 25% respectively 31% of the variance found in the reading achievement. To summarize, when we control for the predictors of reading achievement, the gender differences are considerably reduced. Despite those findings, the gender effect persists.

Moreover, when predicting reading achievement, most of the predictors contribute uniquely. Lastly, VIF and Tolerance values were computed[3].

The VIF values ranged from 1.0 to 1.4, meaning the presence of multicollinearity was minimal. Tolerance values ranged from 0.6 to 0.9, indicating that the multicollinearity was weak. In most cases, the standard errors are the same magnitude as the differences. Therefore, the coefficient and interaction terms are only included to provide a total estimate of the unexplained part.

3.3 Oaxaca-Blinder decomposition Results

Table 7

Summary of Oaxaca-Blinder Decomposition for Norway

	Endowment	**Coefficients**	**Interaction**
	$(X_G-X_B)*\beta_B$	**$(\beta_G-\beta_B)*X_B$**	**$(X_G-X_B)*(\beta_G-\beta_B)$**
Constant		42.27	
joyread	**4.43**	-9.17	6.65
stimread	0.11	-0.29	0.02
workmast	**0.73**	0.33	-0.86
scir	**4.22**	0.07	-0.98
pdir	0.23	0.35	-0.03
readjoyfreq	-1.22	-0.65	0.61
Sum of Component	**8.50 | 18.15%**	**30.64 | 65.44%**	**5.40 | 11.53%**
The Total Sum of Changes, Including the Difference in Constant			**46.82**

Note. Values in bold red represent the significance of the OBD component for each predictor. Coefficients and Interaction terms were not inspected.

3. Table C7 in Appendix C contains the VIF and Tolerance values.

Table C9 and **Table C10** from Appendix C contain more information about the decomposition. **Table 7** and **Table 8** illustrate the summarized Oaxaca-Blinder decomposition results per country.

Before decomposition, t-tests were conducted to test for the gender differences in the means, as suggested by Olsen & Blömeke (2018). It has been found that among our contributing factors (predictors), their means were significantly different, except for one isolated case in Finland respectively two in Norway. Regarding the Oaxaca-Blinder decomposition, all the terms in the formula have to be statistically significant. However, if one of the terms is not significant, then the decomposition for that predictor is not significant. Results from the decomposition indicate that most of the endowments across countries were statistically significant, with a few exceptions. Teachers' stimulation of reading engagement (*stimread*), students' perceived difficulty in reading (*pdir*), and frequency of reading enjoyment (*readjoyfreq*) in Norway. For Finland, teachers' stimulation of reading engagement (*stimread)* and frequency of reading enjoyment (*readjoyfreq)* were insignificant.

In Norway, a total percentage of 18% is explained based on the other characteristics *(endowments),* respectively (8.5 out of 46.8 points), whereas in Finland, the percentage is higher, respectively 23% (15 out of 48.9 points). As a result, the sum of the endowments based on our predictors from the decomposition represents the part of the gap that is reduced if the male group possesses identical characteristics to the female group in each particular country. The *endowments* ranged in Norway from (-1.22 to 4.43), whereas in Finland, the values were higher (-3.13 to 11.96). As previously noticed, the endowments vary and go in both directions. The presence of negative values could mean that changes that positively contribute to the students' reading achievement must first compensate for the negative changes before contributing to explaining the gap (Olsen & Blömeke, 2018, p. 139). This could lead to underestimating the positive predictors, but this could also mean that our model is incomplete

and some other important variables are not included (Olsen & Blömeke, 2018, p. 139). For both countries, *joyread* has the most considerable positive contribution isolated, 9.4% (4.43 points out of 46.82) in Norway, respectively 23.6% (11.56 points out of 48.92) in Finland of the total increase in the reading achievement (*joyread)*. The highest predictor with the negative change is *readjoyfreq* in Norway. Its value of (-1.22) its contribution is negligible since the *endowment* term from this predictor was not significant. For Finland, *pdir* managed to get (-3.13), a moderate negative, statistically significant contribution.

Table 8

Summary of Oaxaca-Blinder decomposition for Finland

	Endowment	**Coefficients**	**Interaction**
	$(X_G-X_B)*\beta_B$	**$(\beta_G-\beta_B)*X_B$**	**$(X_G-X_B)*(\beta_G-\beta_B)$**
Constant		45.59	
joyread	**11.96**	-5.18	6.04
stimread	0.07	0.56	-0.10
workmast	**2.73**	2.74	-1.27
scir	**2.27**	0.53	-0.48
pdir	**-3.13**	-0.65	0.61
readjoyfreq	1.14	0.46	-1.97
Sum of Component	**15.04 | 23.23%**	**31.06 | 47.98%**	**2.82 | 4.36%**
The Total Sum of Changes, Including the Difference in Constant			**48.92**

Note. Values in bold red represent the significance of the OBD term for each predictor. Coefficients and Interaction terms were not inspected.

Chapter 4 Discussion

It has been found that the existence of gender differences in reading persists in Norway and Finland. This study examines how students' characteristics are linked to reading

achievement and how gender differences are accounted for by those characteristics in PISA 2018 reading section in Norway and Finland. This subchapter commences by summarizing and commenting on a) associations between the reading predictors and b) associations between predictors and reading achievement. Then computing the descriptive statistics, applying multilinear regression models, and using Oaxaca-Blinder decomposition.

4.1 Response to the Research Questions

RQ1) The strongest association was discovered to be between students' self-concept in reading (*scir*) and students' perceived difficulty in reading (*pdir)*. This could be due to the nature of those predictors; as known from the conceptual framework, those two are sub-dimensions of the same overarching concept. Next, the predictors of enjoyment of reading (*joyread*) and frequency of reading enjoyment (*readjoyfreq*) were the second-highest correlated. Therefore, it makes sense since both refer to the students' reading enjoyment. Typically, those who enjoy reading more will read for enjoyment more frequently.

RQ2) The predictor that manifested the highest correlation with reading achievement in Norway was students' self-concept in reading (*scir).* In contrast, students' enjoyment of reading (joyread) got the strongest correlation with reading achievement in Finland, in line with the antecedent literature that implies that reading enjoyment and reading achievement are positively correlated (Malanchini et al., 2017). In Norway, the frequency of reading enjoyment (*readjoyfreq*) was the lowest correlation with reading achievement. Teachers' stimulation of reading engagement (*stimread*) was the second least associated predictor with reading achievement. In Finland, it was vice-versa; teachers' stimulation of reading engagement (*stimread*) was the least correlated with the reading achievement, followed by the frequency of reading enjoyment (*readjoyfreq*). However, a weak correlation in Norway and Finland between teachers' stimulation of reading engagement (*stimread)* could be linked to the fact that the

variable was only partly related to the students' reading behavior. Furthermore, our finding aligns with Rojas-Torres and colleagues (2021) study that implies teachers' stimulation of reading engagement is not highly related to students' reading achievement.

RQ3) Out of those two countries, Finland is the country with the highest gender gap in reading achievement. Most of the contributing factors (predictors) means are inclined toward the girls. Essential gender differences have been found in both countries. The most notable ones were found for the enjoyment of reading (*joyread*), self-concept in reading (*scir*), and frequency of reading enjoyment (*readjoyfreq*). This aligns with the previous findings that girls tend to read more for enjoyment and have a higher opinion of themselves regarding their self-concept in reading (Rider & Colmar, 2005). Lastly, girls tend to have a higher frequency of reading enjoyment, assuming that both concepts are interconnected; they read for pleasure, while they get pleasure from reading, resulting in a higher frequency of reading.

RQ4) The results from the pooled multilinear regressions illustrated that when controlling for the set of measures of reading behaviors, the gender difference was substantially reduced in both countries. However, the gender gap was still persistent in both countries. The previous findings show that the gender gap is reducest the most in Finland, followed by Norway. Therefore, it can be concluded that the gender effect was not completely reduced. This could be due to other contributing factors related to the ones that have been used in this study. In addition, based on the regression analysis results, the most substantial contributing factor to reading achievement was students' self-concept in reading (*scir*) in Norway in the pooled sample. Wheres for Finland perceived difficulty in reading (*pdir*) was the most contributing factor to reading achievement.

One to two predictors contributed negatively to the reading achievement across countries students' frequency of reading (*joyreadfreq*) respectively teachers' stimulation of

reading engagement (*stimread*). Only the Finnish male sample from the separate regression analysis did not have regression coefficients that negatively affected reading achievement.

RQ5) This study uses counterfactual questions about the changes in the reading achievement if males had similar coefficients and characteristics as the females. After mean differences in the reading predictors have been decomposed, using the scenario above. By adding variables linked to students' reading behavior in PISA, we can use it to explain some portions of the gap to varying degrees across countries. By taking a deeper look at the decomposition output, it was found that in Norway, students' enjoyment of reading (*joyread)* and students' self-concept in reading (*scir*) were the variables that contributed the most to the gap. In Finland, students' enjoyment of reading (joyread) accounted for a more significant percentage of the gap, followed by students' motivation to master tasks (*workmast*) and students' self-concept in reading (*scir*). Some of the endowments based on predictors negatively contributed to the gap in reading achievement. An insignificant moderate to small negative contribution was found for the frequency of reading enjoyment (*readjoyfreq*) in Norway. A significantly moderate negative contribution to the reading gap was found in Finland for the perceived difficulty in reading (*pdir*).

Interestingly, enjoyment of reading (*joyread*) accounts for almost everything in Finland's endowments term (80%). However, it can be said that this variable is the most important out of all other variables used in our study in Finland. For example, if the Finnish males had the same score in the enjoyment of reading as the Finnish females, their reading achievement would be higher by 11.96 points. The remaining unexplained gap cannot all be explained by the explanatory variables covered in the current study, which indicates there are still a lot of unexplained gaps remaining. Given the strong relationship between the predictors and the extensive literature that those predictions are essential to understanding the gender differences, we must conclude that the explained part was lower than expected. Finally,

numerous other predictors (e.g., reading engagement, intrinsic and extrinsic motivation, self-efficacy) could account for a higher percentage of the reading gap.

4.2 Limitations and Further Research

This study identified a few limitations. Firstly, the assessment design and structure were complex. Secondly, there exists no readily available method to estimate the standard error for decomposition components. We have an abridged solution by ensuring the averages and coefficients differ significantly between the groups. One important limitation is that the frequency of reading enjoyment was based only on one item. We would not expect the association with the reading performance to be very strong. It was necessary to have more predictors in our analysis since one of the decomposition' limitation is that it tends to overestimate the origin of the gaps if the number of variables is not enough (Pelletier & Statistics Canada, 2019). The fact that PISA's questionnaire is a pragmatic and broadly covering instrument. It is not a focused research study of some specific theory. Hence, the inclusion of indicators can not be expected to cover a detailed theoretical framework from educational psychology.

Another limitation discovered was that some variables, such as students' motivation to master tasks (*workmast*) and teachers' stimulation of reading engagement (*stimread),* were not explicitly related to reading. This finding impacted the correlations, regressions, and later the decomposition results. One important finding was that not all the variables in the decomposition contributed positively toward reading achievement. For example, perceived self-concept in reading (*pdir*) in Finland and reading enjoyment frequency (readjoyfreq) in Norway contributed negatively, the *endowments* term of the latter not being significant.

In addition, the data design posed another limitation. This study pooled all the data from three variables *(joyread, stimread, workmast)* across OECD countries. Whereas for the newly developed variables (*scir, pdir, readjoyfreq*), the data were pooled across Norway and Finland.

If all the variables were standardized across OECD, there could have been a slight difference in the results. Therefore, replication of this study's procedures by country is encouraged in future studies for better evidence in the local context. Further comparisons between countries would provide additional insight into the gender differences.

4.3 Conclusions

This research project examined the gender differences in reading in Norway and Finland from PISA 2018. There is more or less a difference between boys and girls in their reading behaviors, and further investigation is needed to determine how different or similar they are. The counterfactual Oaxaca-Blinder decomposition showed that the reading achievement gap would be smaller if boys and girls from Norway and Finland had the same characteristics. Including students' reading behaviors variables such as (*joyread, scir, workmast*) leads to the highest gap contribution in both countries. However, variables such as (*readjoyfreq* and *pdir*) negatively contributed to the reading achievement gap.

The strongest predictor of reading achievement in Finland was students' reading enjoyment (*joyread*), and it was not surprising that it contributed the most to the gap. Whereas for Norway, students' self-concept in reading (*scir*) was the strongest predictor. It was compelling to discover that when it comes to the gap contribution, it was the second contributor after the students' reading enjoyment (*joyread*). Lastly, it can be concluded that the enjoyment of reading (*joyread*) accounted for the most in explaining the gap in both countries.

References

Acar Erdol, T., & Akin-Arikan, Ç. (2022). Gender gap in reading achievement: The mediating role of metacognitive strategies and reading-related attitudes. *Social Psychology of Education*. https://doi.org/10.1007/s11218-022-09692-9

Ammermueller, A. (2007). PISA: What makes the difference?: Explaining the gap in test scores between Finland and Germany. *Empirical Economics*, *33*(2), 263–287. https://doi.org/10.1007/s00181-006-0102-5

Bandura, A. (1986). The Explanatory and Predictive Scope of Self-Efficacy Theory. *Journal of Social and Clinical Psychology*, *4*(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359

Bangdiwala, S. I. (2018). Regression: Multiple linear. *International Journal of Injury Control and Safety Promotion*, *25*(2), 232–236. https://doi.org/10.1080/17457300.2018.1452336

Barrera-Osorio, F., Garcia-Moreno, V., Patrinos, H. A., & Porta, E. (2011). *Using the Oaxaca-Blinder decomposition technique to analyze learning outcomes changes over time: An application to Indonesia's results in PISA mathematics*. The World Bank. https://doi.org/10.1596/1813-9450-5584

Botezat, A. (2012). Decomposing the gap in school achievement between Finland and Romania: Some methodological aspects. *Annals of Faculty of Economics, 1*(2), 165–171.

Brumfit, C. (1993). Review of The Braid of Literature: Children's Worlds of Reading [Review of *Review of The Braid of Literature: Children's Worlds of Reading*, by S. A. Wolff & S. B. Heath]. *British Journal of Educational Studies*, *41*(3), 307–309. https://doi.org/10.2307/3122294

Cambria, J., & Guthrie, J. T. (2010). Motivating and engaging students in reading. *New*

England Reading Association Journal, *46*(1), 16–29.

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, *87*(1), 154–167. https://doi.org/10.1037/0022-0663.87.1.154

Cheema, J. R. (2018). Adolescents' enjoyment of reading as a predictor of reading achievement: New evidence from a cross-country survey: Enjoyment of Reading and Reading Achievement. *Journal of Research in Reading*, *41*, S149–S162. https://doi.org/10.1111/1467-9817.12257

Chyung, S. Y. Y., Barkin, J. R., & Shamsy, J. A. (2018). Evidence-Based Survey Design: The Use of Negatively Worded Items in Surveys. *Performance Improvement*, *57*(3), 16–25. https://doi.org/10.1002/pfi.21749

Clark, C. & Rumbold, K. (2006). *Reading for pleasure: a Research Overview*. London: National Literacy Trust.

Combs, A. J. (1962). *Perceiving Behaving Becoming. A New Focus for Education*. Association for Supervision and Curriculum Development, Suite 1100. http://eric.ed.gov/ERICWebPortal/detail?accno=ED096575

Crato, N. (Ed.). (2021). *Improving a Country's Education: PISA 2018 Results in 10 Countries*. Springer International Publishing. https://doi.org/10.1007/978-3-030-59031-4

De Naeghel, J., Van Keer, H., Vansteenkiste, M., & Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: A self-determination theory perspective. *Journal of Educational Psychology*, *104*(4), 1006. https://doi.org/10.1037/a0027800

Deci, E. L., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology*, *40*(1), 1–10. https://doi.org/10.1037/0022-3514.40.1.1

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US. https://doi.org/10.1007/978-1-4899-2271-7

Deci, E. L., & Ryan, R. M. (2004). *Handbook of Self-determination Research*. University Rochester Press.

Diseth, Å., Danielsen, A. G., & Samdal, O. (2012). A path analysis of basic need support, self-efficacy, achievement goals, life satisfaction and academic achievement level among secondary school students. *Educational Psychology*, *32*(3), 335–354. https://doi.org/10.1080/01443410.2012.657159

Elder, T. E., Goddeeris, J., & Haider, S. J. (2009). A Deadly Disparity: A Comprehensive Assessment of the Black-White Infant Mortality Gap. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1493334

Federici, R., & Skaalvik, E. (2013). Students' Perceptions of Emotional and Instrumental Teacher Support: Relations with Motivational and Emotional Responses. *International Education Studies*, *7*(1), p21. https://doi.org/10.5539/ies.v7n1p21

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics, Vol. 4B.* (pp. 1-97). North-Holland.

Frønes, T. S., Pettersen, A., Radišić, J., & Buchholtz, N. (2020). *Equity, Equality and Diversity in the Nordic Model of Education—Contributions from Large-Scale Studies.* In *Diversity in the Nordic Model of Education* (pp. 1-10). (1st ed. 2020.). Springer International Publishing : Imprint: Springer.

Gevrek, Z. E., & Seiberlich, R. R. (2014). Semiparametric decomposition of the gender achievement gap: An application for Turkey. *Labour Economics*, *31*, 27–44. https://doi.org/10.1016/j.labeco.2014.08.002

Giménez, G., & Aristizábal, G. C. (2017). Por qué los estudiantes de colegios públicos y privados de Costa Rica obtienen distintos resultados académicos? *Perfiles latinoamericanos : revista de la Sede Académica de México de la Facultad Latinoamericana de Ciencias Sociales*, *25*(49), 195–222. https://doi.org/10.18504/pl2549-009-2017

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, Gender, and Math. *Science*, *320*(5880), 1164–1165. https://doi.org/10.1126/science.1154094

Guthrie, J. T., Hoa, A. L. W., Wigfield, A., Tonks, S. M., Humenick, N. M., & Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology*, *32*(3), 282–313. https://doi.org/10.1016/j.cedpsych.2006.05.004

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hlavac, M. (2014). oaxaca: Blinder-Oaxaca Decomposition in R. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2528391

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jabbar, A., & Warraich, N. F. (2022). Gender differences in leisure reading habits: A systematic review of literature. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-12-2020-0200

Jann, B. (2008). The Blinder–Oaxaca Decomposition for Linear Regression Models. *The Stata Journal*, *8*(4), 453–479. https://doi.org/10.1177/1536867X0800800401

Jeno, L. M., Raaheim, A., Kristensen, S. M., Kristensen, K. D., Hole, T. N., Haugland, M. J., & Mæland, S. (2017). The Relative Effect of Team-Based Learning on Motivation and Learning: A Self-Determination Theory Perspective. *CBE—Life Sciences Education*, *16*(4), ar59. https://doi.org/10.1187/cbe.17-03-0055

Kim, J.-O., & Curry, J. (1977). The Treatment of Missing Data in Multivariate Analysis. *Sociological Methods & Research*, *6*(2), 215–240. https://doi.org/10.1177/004912417700600206

Kröger, H., & Hartmann, J. (2021). Extending the Kitagawa–Oaxaca–Blinder decomposition approach to panel data. *The Stata Journal*, *21*(2), 360–410. https://doi.org/10.1177/1536867X211025800

Larsen, J. E., Schulte, B., & Thue, F. W. (2021). *Schoolteachers and the Nordic Model: Comparative and Historical Perspectives* (1st ed.). Routledge. https://doi.org/10.4324/9781003082514

Loveless, T. (2015). The 2015 Brown Center Report on American Education: How Well Are American Students Learning? With Sections on the Gender Gap in Reading, Effects of the Common Core, and Student Engagement. Volume 3, Number 4. In *Brookings Institution*. Brookings Institution.

Maitner, A. T., & Summerville, A. (2022). "What was meant to be" versus "what might have been": Effects of culture and control on counterfactual thinking. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000295

Malanchini, M., Wang, Z., Voronin, I., Schenker, V. J., Plomin, R., Petrill, S. A., & Kovas, Y. (2017). Reading self-perceived ability, enjoyment and achievement: A genetically

informative study of their reciprocal links over time. *Developmental Psychology*, *53*(4), 698–712. https://doi.org/10.1037/dev0000209

Marsh, H. W. (1986). Negative Item Bias in Ratings Scales for Preadolescent Children: A Cognitive-Developmental Phenomenon. *Developmental Psychology*, *22*(1), 37–49. https://doi.org/10.1037/0012-1649.22.1.37

Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multidimensional self-concepts: Relations with sex and academic achievement. *Journal of Educational Psychology*, *77*(5), 581–596. https://doi.org/10.1037/0022-0663.77.5.581

Marshall, H. H. (1987). Motivational Strategies of Three Fifth-Grade Teachers. *Elementary School Journal*, *88*(2), 135–150.

Mikheeva, E., & Meyer, S. (2019). *Help manual for the IEA IDB analyzer (Version 4.0)*. 274.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. TIMSS & PIRLS International Study Center, Boston College.

Munita, F. (2016). Prácticas didácticas, creencias y hábitos lectores del profesor en una escuela exitosa en la promoción lectora. [Didactic practices, credentials and professors' lectures by a school exit in the reading promotion].*Ocnos. Revista de estudios sobre lectura*, *15*(2), 77–97. https://doi.org/10.18239/ocnos_2016.15.2.1140

National Center for Education Statistics. (2019).*The nations report card*: 2019 Mathematics and reading assessments. U.S. Department of Edu-cation, National Center for Education Statistics, Office of Educational Research and Improvement.

Neimeyer, R. A., Pitcho-Prelorentzos, S., & Mahat-Shamir, M. (2021). "If only…": Counterfactual thinking in bereavement. *Death Studies*, *45*(9), 692–701. https://doi.org/10.1080/07481187.2019.1679959

Ng, C., & Bartlett, B. (Eds.). (2017). *Improving Reading and Reading Engagement in the 21st Century*. Springer Singapore. https://doi.org/10.1007/978-981-10-4331-4

Nicholls, J. G., & Miller, A. T. (1983). The Differentiation of the Concepts of Difficulty and Ability. *Child Development*, *54*(4), 951–959. https://doi.org/10.1111/j.1467-8624.1983.tb00516.x

Nieto, S., & Ramos, R. (2015). Educational outcomes and socioeconomic status: A decomposition analysis for middle-income countries. *Prospects (Paris)*, *45*(3), 325–343. https://doi.org/10.1007/s11125-015-9357-y

Ning, B., Van Damme, J., Van Den Noortgate, W., Gielen, S., Bellens, K., Dupriez, V., & Dumay, X. (2016). Regional inequality in reading performance: An exploration in Belgium. *School Effectiveness and School Improvement*, *27*(4), 642–668. https://doi.org/10.1080/09243453.2016.1202842

OECD. (2010). The Reading and Learning Habits of 15-Year-Olds. In OECD, *PISA 2009 Results: Learning to Learn* (pp. 59–83). OECD. https://doi.org/10.1787/9789264083943-6-en

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD. https://doi.org/10.1787/9789264266490-en

OECD. (2018). *PISA 2018 Technical Report.* https://www.oecd.org/pisa/data/pisa2018technicalreport/

OECD. (2019a). *PISA 2018 Results (Volume I): What Students Know and Can Do*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-i_5f07c754-en

OECD. (2019b). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-ii_b5fd1b8f-en

OECD. (2019c). *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-iii_acd78851-en

OECD. (2019d). *PISA 2018 Assessment and Analytical Framework*. OECD. https://doi.org/10.1787/b25efab8-en

OECD. (2019e). *PISA 2018 Reading Framework*. 21–71. https://doi.org/10.1787/5c07e4f1-en

OECD. (2019f). *PISA 2018 Background questionnaires* (pp. 1–83). OECD. https://doi.org/10.1787/67e1518f-en

OECD. (2022). *PISA 2000-2018, Reading, Mathematics, and Science Assessments*. Retrieved 20 January, 2022 from https://pisadataexplorer.oecd.org/ide/idepisa/

Olastuen, C. (2018). *Jeg elsker å lese! ...eller? En kvalitativ studie av lesemotivasjon på ungdomstrinnet*.[I love reading! ...or? A qualitative study of reading motivation at the lower secondary level]. http://urn.nb.no/URN:NBN:no-69941

Pelletier, R., Patterson, M., & Moyser, M. (2019). The gender wage gap in Canada: 1998 to 2018. *Labour Statistics*: Research Papers. Statistics Canada, Catalogue no. *75-004-M*.

Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, *33*(4), 335–355. https://doi.org/10.1111/j.1467-9817.2009.01418.x

Reeve, J. (2009). Why Teachers Adopt a Controlling Motivating Style Toward Students and How They Can Become More Autonomy Supportive. *Educational Psychologist*, *44*(3), 159–175. https://doi.org/10.1080/00461520903028990

Reilly, D. (2012). Gender, Culture, and Sex-Typed Cognitive Abilities. *PLoS ONE*, *7*(7), e39904. https://doi.org/10.1371/journal.pone.0039904

Reynolds, R., Arnone, M., & Marshall, T. (2009). Perceived competence and reading enjoyment as contributors to information skills and digital technology knowledge: Perceived Competence and Reading Enjoyment as Contributors to Information Skills and Digital Technology Knowledge. *Proceedings of the American Society for Information Science and Technology*, *46*(1), 1–26. https://doi.org/10.1002/meet.2009.1450460235

Rider, N., & Colmar, S. (2005). Reading achievement and reading self-concept in year 3 children [Electronic version]. *Australian Association for Research in Education*. Retrieved February 8, 2022 from https://www.aare.edu.au/data/publications/2005/col05347.pdf

Roe, A., & Jensen, R. E. (2017). Hva kjennetegner de svakeste elevenes lesekompetanse, lesevaner og strategikunnskap når de går ut av grunnskolen? [What are the characteristics of the weakest students' reading skills, reading habits and strategic knowledge when they leave primary school?]. *Nordic Journal of Literacy Research*, *3*(2), Article 2. https://doi.org/10.23865/njlr.v3.882

Roese, N. J., Smallman, R., & Epstude, K. (2017). Do episodic counterfactual thoughts focus on controllable action?: The role of self-initiation. *Journal of Experimental Social Psychology*, *73*, 14–23. https://doi.org/10.1016/j.jesp.2017.05.006

Rojas-Torres, L., Ordóñez, G., & Calvo, K. (2021). Teacher and Student Practices Associated with Performance in the PISA Reading Literacy Evaluation. *Frontiers in Education*, *6*. https://www.frontiersin.org/article/10.3389/feduc.2021.658973

Olsen, R. V., & Blömeke, S. (2018). *Hva forklarer endringer i elevenes matematikkprestasjoner over tid?* [What explains the change in students' mathematics performance over time?]. In J. K. Björnsson & R. V. Olsen (Eds.), *Tjue år med TIMSS og PISA i Norge: Trender og nye analyser* [Twenty years with TIMSS And PISA in Norway: Trends and new analysis] (pp. 128–149). Oslo: Universitetsforlaget.

Rosén, M., Steinmann, I., & Wernersson, I. (2022). Gender Differences in School Achievement. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International Handbook of Comparative Large-Scale Studies in Education* (pp. 1–48). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_46-1

Ruddell, R. B. (1995). Those Influential Literacy Teachers: Meaning Negotiators and Motivation Builders. *The Reading Teacher*, *48*(6), 454–463.

Rust, K., & Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, *5*(3), 283–310. https://doi.org/10.1177/096228029600500305

Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, *42*, 95–103. https://doi.org/10.1016/j.learninstruc.2016.01.004

Ryan, R. M., & Deci, E. L. (2000a). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2000b). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics classrooms? *Journal of School Psychology*, *50*(2), 235–255. https://doi.org/10.1016/j.jsp.2011.10.005

San Román, A. G. de, & de La Rica, S. (2015). Gender Gaps in Spain: The Role of Children in Career Development. In *Unexplored Dimensions of Discrimination*. University Press. https://doi.org/10.1093/acprof:oso/9780198729853.003.0004

Sari, I. M. (2020). *Decomposing Gender Differences in Reading: Evidence from Nordic PIRLS 2016*. https://www.duo.uio.no/handle/10852/84706

Scholes, L. (2019). Differences in attitudes towards reading and other school-related activities among boys and girls. *Journal of Research in Reading*, *42*(3–4), 485–503. https://doi.org/10.1111/1467-9817.12279

Schulz-Heidorf, K., & Støle, H. (2018). Gender differences in Norwegian PIRLS 2016 and ePIRLS 2016 results at test mode, text and item format level. *Nordic Journal of Literacy Research*, *4*(1), 167-183. Article 1. https://doi.org/10.23865/njlr.v4.1270

Schunk, D. H., Meece, J. R., & Pintrich, P. R. (2012). Motivation in education: Theory, research, and applications. *Upper Saddle River*: Pearson Higher Ed.

Sen, B. (2014). Using the oaxaca-blinder decomposition as an empirical tool to analyze racial disparities in obesity: Oaxaca-Blinder Decomposition and Obesity Disparities. *Obesity (Silver Spring, Md.)*, *22*(7), 1750–1755. https://doi.org/10.1002/oby.20755

Shiel, G., & Cosgrove, J. (2002). International Perspectives on Literacy: International Assessments of Reading Literacy. *The Reading Teacher*, *55*(7), 690–692.

The United Nations Educational, Scientific and Cultural Organization. (2011). *Armed conflict and education*; *EFA global monitoring report*. Retrieved 4 February 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000190743

Torppa, M., Eklund, K., Sulkunen, S., Niemi, P., & Ahonen, T. (2018). Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure reading and homework activity. *Journal of Research in Reading*, *41*(1), 122–139. https://doi.org/10.1111/1467-9817.12103

Tunstall, P., & Gsipps, C. (1996). Teacher Feedback to Young Children in Formative Assessment: A typology. *British Educational Research Journal*, *22*(4), 389–404. https://doi.org/10.1080/0141192960220402

van Langen, A., Bosker, R., & Dekkers, H. (2006). Exploring cross-national differences in gender gaps in education. *Educational Research and Evaluation*, *12*(2), 155–177. https://doi.org/10.1080/13803610600587016

Von Davier, M., Gonzalez, E., & Mislevy. R. *(2009). What are plausible values and why are they useful. IERI monograph series, 2, 9–36.*

Ward, M. (2020). *PISA for Development: Out-of-school assessment: Results in Focus*. OECD. https://doi.org/10.1787/491fb74a-en

Werlich, E. (1976). *A text grammar of English*. Quelle und Meyer.

Wray, D., Medwell, J., Fox, R., & Poulson, L. (2000). The Teaching Practices of Effective Teachers of Literacy. *Educational Review*, *52*(1), 75–84. https://doi.org/10.1080/00131910097432

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, *31*(2–3), 114–128. https://doi.org/10.1016/j.stueduc.2005.05.005

Yu, R., & Singh, K. (2018). Teacher support, instructional practices, student motivation, and mathematics achievement in high school. *The Journal of Educational Research*, *111*(1), 81–94. https://doi.org/10.1080/00220671.2016.1204260

Zasacka, Z., & Bulkowski, K. (2017). Reading engagement and school achievement of lower secondary school students. *Edukacja*, *141*(2). https://doi.org/10.24131/3724.170205

Zehner, F., Goldhammer, F., & Sälzer, C. (2018). Automatically analyzing text responses for exploring gender-specific cognitions in PISA reading. *Large-Scale Assessments in Education*, *6*(7), 26-6:7<26. https://doi.org/10.1186/s40536-018-0060-3

Zimmerman, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. *Handbook of Self-Regulation*, 13–39. https://doi.org/10.1016/B978-012109890-2/50031-7

Appendices

Appendix A Ethical Approval and GDPR

According to the (NSD) Norwegian Centre for Research Data notification procedure, this project fulfills its duty imposed by the EEA's general data protection regulation (GDPR). There is no way to link the PISA 2018 data back to any individual participant, as it is composed of only aggregated and depersonalized data. The author did not collect or use identifiable personal data during the research process. Therefore, this project is not subject to the GDPR notification. Thus, this project adheres to the University of Oslo guidelines regarding ethical approval.

Appendix B: Data Management & Analysis Code

B.1 Computation of the New Indices "SCIR," "READJOYFREQ," and "PDIR"

```
* Encoding: UTF-8.

DATASET ACTIVATE DataSet1.
COMPUTE SCIR=MEAN(ST161Q01HA, ST161Q02HA, ST161Q03HA).
EXECUTE.

WEIGHT BY W_FSTUWT.

DESCRIPTIVES VARIABLES=SCIR
  /SAVE
  /STATISTICS=MEAN STDDEV MIN MAX.

RECODE ST175Q01IA (1=0) (2=1) (3=2) (4=3) (5=4) (SYSMIS=SYSMIS).
EXECUTE.

COMPUTE READJOYFREQ=MEAN(ST175Q01IA).
EXECUTE.

WEIGHT BY W_FSTUWT.

DESCRIPTIVES VARIABLES=READJOYFREQ
/SAVE
/STATISTICS=MEAN STDDEV MIN MAX.

RECODE ST161Q06HA (4=1) (3=2) (2=3) (1=4) (SYSMIS=SYSMIS) INTO
ST161Q06R.
EXECUTE.

 RECODE ST161Q07HA (4=1) (3=2) (2=3) (1=4) (SYSMIS=SYSMIS) INTO
ST161Q07R.
 EXECUTE.

RECODE ST161Q08HA (4=1) (3=2) (2=3) (1=4) (SYSMIS=SYSMIS) INTO
ST161Q08R.
EXECUTE.

COMPUTE PDIR=MEAN(ST161Q06R, ST161Q07R, ST161Q08R).
EXECUTE.

WEIGHT BY W_FSTUWT.

DESCRIPTIVES VARIABLES=PDIR
  /SAVE
  /STATISTICS=MEAN STDDEV MIN MAX.
```

B.2 Associations Among the Variables

```
* Encoding: UTF-8.

* Script created using the IEA IDB Analyzer (Version 5.0.5).

* Created on 2/23/2022 at 12:52 AM.

* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
"C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
JB_Cor.ieasps".

JB_Cor infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/

        cvar=CNTRYID /

        xvar=JOYREAD STIMREAD WORKMAST SCIR PDIR READJOYFREQ /

        wgt=W_FSTUWT/

        rwgt=W_FSTURWT/

        kfac=0.5/

        jk2type=/

        stratvar=/

        nrwgt=80/

        nomiss=Y/

        strctry = N/

        method=BRR/

        pairwise=Y/

        viewcod=N/

        ndec=2/

        clean = Y/

        selcrit = /

        selvar = /

        intavg = Y/

        outdir="C:\Users\Iulian\Desktop\PISA THESIS"/

        outfile="CORRELATIONPAIRWISE".
```

B.3 Associations Between Student's Attributes and the Reading Achievement

```
* Script created using the IEA IDB Analyzer (Version 5.0.5).

* Created on 3/15/2022 at 8:47 PM.
```

```
* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
"C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
JB_CorP.ieasps".

JB_CorP         infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/

        cvar=CNTRYID /

        xvar0=JOYREAD PDIR SCIR READJOYFREQ STIMREAD WORKMAST /

        PVRoots=PV /

        PVTails=READ /

        npv=10/

        wgt=W_FSTUWT/

        nrwgt=80 /

        rwgt=W_FSTURWT/

        jkz=/

        jkr=/

        jk2type=/

        stratvar=/

        nomiss=Y/

        strctry = N/

        method=BRR/

        pairwise=N/

        kfac=0.5/

        shrtcut=N/

        viewcod=N/

        ndec=2/

        clean = Y/

        selcrit = /

        selvar = /

        intavg = Y/

        outdir="C:\Users\Iulian\Desktop\PISA THESIS"/

        outfile="PVANDCORR".
```

B.4 Reading Achievement Average Based on Gender

```
* Encoding: windows-1252.
* Script created using the IEA IDB Analyzer (Version 5.0.5).
* Created on 2/23/2022 at 1:23 AM.
* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
"C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
JB_PV.ieasps".

JB_PV           infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/
        cvar=CNTRYID ST004D01T /
        almvars=/
        rootpv=PV /
        tailpv=READ /
        npv=10/
        wgt=W_FSTUWT/
        nrwgt=80 /
        rwgt=W_FSTURWT/
        jkz=/
        jkr=/
        jk2type=/
        stratvar=/
        nomiss=Y/
        method=BRR/
        kfac=0.5/
        shrtcut=N/
        viewcod=N/
        ndec=2/
        clean = Y/
        strctry = N/
        intavg = Y/
        graphs=Y/
        selcrit = /
        selvar = /
        outdir="C:\Users\Iulian\Desktop\PISA THESIS"/
        outfile="meanreadingacchivbygender".
```

B.5 The Analysis of the Mean Difference

```
* Encoding: windows-1252.
* Script created using the IEA IDB Analyzer (Version 5.0.5).
* Created on 3/4/2022 at 7:25 PM.
* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
"C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
JB_Gen.ieasps".
```

```
JB_gen          infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/
          cvar=CNTRYID ST004D01T /
          almvars=/
          xvar=JOYREAD STIMREAD WORKMAST SCIR PDIR READJOYFREQ /
          wgt=W_FSTUWT/
          rwgt=W_FSTURWT/
          kfac=0.5/
          jk2type=/
          stratvar=/
          nrwgt=80/
          nomiss=Y/
          method=BRR/
          viewcod=N/
          ndec=2/
          clean = Y/
          strctry = N/
          intavg = Y/
          graphs=Y/
          selcrit = /
          selvar = /
          outdir="C:\Users\Iulian\Desktop\PISA THESIS"/
          outfile="MeanDifferenceAnalysis".
```

B.6 Analysis of Pooled Regressions Using Gender as a Dummy Variable

```
* Encoding: UTF-8.
* Script created using the IEA IDB Analyzer (Version 5.0.5).
* Created on 2/23/2022 at 1:51 AM.
* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
  "C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
  JB_RegGP.ieasps".

JB_RegGP        infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/
   cvar=CNTRYID /
   convar=JOYREAD STIMREAD WORKMAST SCIR PDIR READJOYFREQ  /
   catvar=ST004D01T /
   codings=D/
   refcats=1/
   ncats=2/
   PVRoots=/
   PVTails=/
   dvar0=/
   rootpv=PV /
   tailpv=READ /
   npv=10/
   wgt=W_FSTUWT/
   nrwgt=80 /
   rwgt=W_FSTURWT/
```

```
jkz=/
jkr=/
jk2type=/
stratvar=/
nomiss=Y/
method=BRR/
missing=listwise/
kfac=0.5/
shrtcut=N/
viewcod=N/
ndec=2/
clean = Y/
strctry = N/
viewprgs=Y/
viewlbl=Y/
qcstats=Y/
newout=Y/
intavg = Y/
selcrit = /
selvar = /
outdir="C:\Users\Iulian\Desktop\PISA THESIS"/
outfile="PRGenderDummyVariable".
```

B.7 Gender-Specific Regression Models

```
* Encoding: UTF-8.

* Script created using the IEA IDB Analyzer (Version 5.0.5).

* Created on 2/23/2022 at 1:51 AM.

* Press Ctrl+A followed by Ctrl+R to submit this analysis.

include file =
"C:\Users\Iulian\AppData\Roaming\IEA\IDBAnalyzerV5\bin\Data\Templates\SPSS_Macros\
JB_RegGP.ieasps".

JB_RegGP      infile="C:\Users\Iulian\Desktop\PISA THESIS\PISA_2018.sav"/

        cvar=CNTRYID ST004D01T /

        convar=JOYREAD STIMREAD WORKMAST SCIR PDIR READJOYFREQ /

        catvar=/

        codings=D/

        refcats=1/

        ncats=2/

        PVRoots=/

        PVTails=/

        dvar0=/
```

```
rootpv=PV /
tailpv=READ /
npv=10/
wgt=W_FSTUWT/
nrwgt=80 /
rwgt=W_FSTURWT/
jkz=/
jkr=/
jk2type=/
stratvar=/
nomiss=Y/
method=BRR/
missing=pairwise/
kfac=0.5/
shrtcut=N/
viewcod=N/
ndec=2/
clean = Y/
strctry = N/
viewprgs=Y/
viewlbl=Y/
qcstats=Y/
newout=Y/
intavg = Y/
selcrit = /
selvar = /
outdir="C:\Users\Iulian\Desktop\PISA THESIS"/
outfile="SEPARATEBYGENDER".
```

B.8 R Script of CFA

```
library(readxl)
library(lavaan)
```

```
library(semPlot)
library(psych)
library(haven)
setwd("C:/Users/Iulian/Desktop/PISA THESIS")
PISA2018_Copy <- read_sav("C:/Users/Iulian/Desktop/PISA THESIS/PISA2018 - Copy.sav")
SCIR<-"SCIR =~ ST161Q01HA + ST161Q02HA + ST161Q03HA + ST161Q06R + ST161Q07R + ST161Q08R"
SCIR1 <- cfa(SCIR,
        data = PISA2018_Copy,
        std.lv = TRUE, estimator = "ML")
summary(SCIR1, standardized = TRUE, rsquare = TRUE, fit.measures = TRUE)
SCIRANDPDIR<-"SCIR =~ ST161Q01HA + ST161Q02HA + ST161Q03HA
      PDIR =~ ST161Q06R + ST161Q07R + ST161Q08R"
SCIRANDPDIR1 <- cfa(SCIRANDPDIR,
        data = PISA2018_Copy,
        std.lv = TRUE, estimator = "ML")
summary(SCIRANDPDIR1, standardized = TRUE, rsquare = TRUE, fit.measures = TRUE)
semPaths(SCIR1,
      whatLabels="std", intercepts=FALSE, style="lisrel",
      nCharNodes=0,
      nCharEdges=0,
      rotation = 4,
      edge.label.position=0.5,
      edge.label.cex = 0.8,
      sizeMan=4.0,
      curveAdjacent = TRUE,title=TRUE, layout="tree2",curvePivot=TRUE,color = "lightblue")
semPaths(SCIRANDNPDIR1,
      whatLabels="std", intercepts=FALSE, style="lisrel",
      nCharNodes=0,
      nCharEdges=0,
```

```
        rotation = 4,
        edge.label.position=0.5,
        edge.label.cex = 0.8,
        sizeMan=4.0,
        curveAdjacent = TRUE,title=TRUE, layout="tree2",curvePivot=TRUE,color =
"lightblue")
anova(SCIR1,SCIRANDPDIR1)
write_sav(PISA2018_Copy, "PISA_2018.sav")
```

Appendix C: SupplementaryTable C

Table C1

Student Sample Size and Gender

Country	N	Gender	
		Male	Female
Norway	5813	2933	2880
Finland	5649	2877	2722

Table C2

The Missing Values in Norway and Finland for each Variable

Variable	Norway			Finland		
	N	NA	%	N	NA	%
gender	5813	0	0	5649	0	0
joyread	5521	561	10.1	5489	314	5.7
stimread	5467	656	12	5419	450	8.3
workmast	5347	855	16	5269	727	13.8
scir	5501	594	10.8	5453	387	7.1
pdir	5493	608	11.1	5450	397	7.3
readjoyfreq	5469	654	12	5462	310	5.7

Note. NA= missing values. The values in the "%" columns were rounded to the nearest .1

Table C3

CFA Results for the Single-Factor Model

Variable	Items	Factor loading	CFI	TLI	RMSEA	SRMR
ST161Q01HA	I am a good reader	0.83	0.79	0.65	0.24	0.11
ST161Q02HA	I am able to understand difficult texts.	0.80				
ST161Q03HA	I read fluently.	0.79				
ST161Q06R	I have always had difficulty with reading.	0.43				
ST161Q07R	I have to read a text several times before completely understanding it.	0.47				
ST161Q08R	I find it difficult to answer questions about a text.	0.48				

Note. The factor loadings illustrated are standardized.

Table C4

CFA Results for the Two-Factor Model

Variable	Factor	Items	Factor loading	CFI	TLI	RMSEA	SRMR
ST161Q01HA	1	I am a good reader	0.85	0.97	0.95	0.08	0.03
ST161Q02HA	1	I am able to understand difficult texts.	0.81				
ST161Q03HA	1	I read fluently.	0.80				
ST161Q06R	2	I have always had difficulty with reading.	0.59				
ST161Q07R	2	I have to read a text several times before completely understanding it.	0.79				

ST161Q08R	2	I find it difficult to answer questions about a text.	0.76

Note. The factor loadings illustrated are standardized.

Table C5

Output of ANOVA

Name	Df	AIC	BIC	Chisq	Chisq diff	Df diff	Pr(>Chisq)
SCIRANDPDIR1	8	130026	130121	641.5			
SCIR1	9	134938	135025	5555.4	4913.9	1	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Note. SCIRANDPDIR1 represents the two-factor model and SCIR1 the single-factor model.

Table C6

Multiple Regression Results with Pooled Sample and Separated by Gender

	Norway			Finland		
	Pooled	Girls	Boys	Pooled	Girls	Boys
	β (s.e)	β (s.e)	β (s.e)	β (s.e)	β (s.e)	β (s.e)
Intercept	**530.47** (2.85)	**527.53** (2.62)	**485.26** (2.78)	**550.54** (1.93)	**546.22** (2.08)	**510.63** (3.24)
Male	**-27.19** (2.80)			**-30.73** (2.72)		
joyread	**14.36** (1.17)	**19.11** (1.66)	**7.64** (1.92)	**19.36** (1.49)	**25.73** (2.01)	**17.09** (2.10)
stimread	**5.50** (1.67)	**6.91** (2.01)	**5.56** (2.17)	-0.25 (1.22)	-0.78 (1.75)	1.81 (1.93)
workmast	2.09 (1.43)	-1.03 (1.94)	**5.65** (1.90)	**11.07** (1.34)	**7.70** (1.92)	**14.39** (2.26)
scir	**26.25** (1.75)	**23.10** (2.65)	**30.15** (2.03)	**19.42** (1.62)	**17.87** (2.60)	**22.73** (2.42)
pdir	**19.57** (1.40)	**19.45** (2.15)	**23.03** (2.30)	**17.08** (1.55)	**15.76** (2.49)	**19.62** (2.09)
readjoyfreq	-4.39 (1.60)	-3.79 (1.84)	-7.65 (2.59)	0.77 (1.23)	-2.45 (1.49)	3.37 (2.38)

R-Square	0.25 (0.01)	0.25 (0.02)	0.21 (0.02)	0.31 (0.01)	0.29 (0.02)	0.25 (0.02)

Note. Significant coefficients are marked in bold.

Table C7

Collinearity Diagnostics for the Pooled Sample

Coefficients[a]

Model		Collinearity Statistics	
	Sig.	Tolerance	VIF
(Constant)	.00		
Joy/Like reading (JOYREAD)	.00	.75	1.31
Teacher's stimulation of reading engagement perceived by the student (STIMREAD)	<.001	.94	1.06
Motivation to master tasks (WORKMAST)	<.001	.89	1.11
Self-concept in reading (SCIR)	.00	.68	1.46
Perception of difficulty in reading (PDIR)	.00	.77	1.29
Frequency of reading enjoyment (READJOYFREQ)	.93	.84	1.17

Note. a. Dependent variable: plausible value 1 in reading.

Table C8

Inter-Item Correlation Matrix

Item	ST161Q01HA	ST161Q02HA	ST161Q03HA	ST161Q06R	ST161Q07R	ST161Q08R
ST161Q01HA	—	.69	.70	.32	.31	.33
ST161Q02HA	.69	—	.64	.26	.38	.39
ST161Q03HA	.70	.64	—	.35	.32	.32
ST161Q06R	.32	.26	.35	—	.47	.42
ST161Q07R	.31	.38	.32	.47	—	.60
ST161Q08R	.33	.39	.32	.42	.60	—

NORWAY

Table C9

Decomposition Results for Norway

	Regression Analysis PISA 2018				Oaxaca-Blinder decomposition PISA 2018		
	Girls		Boys		**Endowment (% of the total gap)**	**Coefficients (% of the total gap)**	**Interaction (% of the total gap)**
	$\mathbf{X_G}$ (s.e)	$\boldsymbol{\beta_G}$ (s.e)	$\mathbf{X_B}$ (s.e)	$\boldsymbol{\beta_B}$ (s.e)	$\mathbf{(X_G-X_B)*\beta_B}$	$\mathbf{(\beta_G-\beta_B)*X_B}$	$\mathbf{(X_G-X_B)*(\beta_G-\beta_B)}$
Intercept		527.53 (2.62)		485.26 (2.78)	-	42.27	-
joyread	-0.22 (0.02)	19.11 (1.66)	-0.80 (0.02)	7.64 (1.92)	**4.43**	-9.17	6.65
stimread	-0.20 (0.02)	6.91 (2.01)	-0.22 (0.02)	5.56 (2.17)	0.11	-0.29	0.02
workmast	0.08 (0.02)	-1.03 (1.94)	-0.05 (0.02)	5.65 (1.90)	**0.73**	0.33	-0.86
scir	0.13 (0.02)	23.10 (2.65)	-0.01 (0.02)	30.15 (2.03)	**4.22**	0.07	-0.98
pdir	-0.09 (0.02)	19.45 (2.15)	-0.10 (0.02)	23.03 (2.30)	0.23	0.35	-0.03
readjoyfreq	-0.01 (0.03)	-3.79 (1.84)	-0.17 (0.01)	-7.65 (2.59)	-1.22	-0.65	0.61
R-Square		0.25 (0.02)		0.21 (0.02)	**8.50 | 18.15%**	**30.64 | 65.44%**	**5.40 | 11.53%**
					The Total Sum of Changes, Including the Difference in Constant		**46.82**

Note. Values in bold red represent the significance of the OBD component for each predictor. Coefficients and Interaction terms were not inspected.

FINLAND

Table C10

Decomposition Results for Finland

	Regression Analysis PISA 2018				Oaxaca-Blinder decomposition PISA 2018		
	Girls		Boys		**Endowment (% of the total gap)**	**Coefficients (% of the total gap)**	**Interaction (% of the total gap)**
	$\mathbf{X_G}$ (s.e)	$\boldsymbol{\beta_G}$ (s.e)	$\mathbf{X_B}$ (s.e)	$\boldsymbol{\beta_B}$ (s.e)	$\mathbf{(X_G-X_B)*\beta_B}$	$\mathbf{(\beta_G-\beta_B)*X_B}$	$\mathbf{(X_G-X_B)*(\beta_G-\beta_B)}$
Intercept		546.22 (2.08)		510.63 (3.24)	-	45.59	-
joyread	0.10 (0.01)	25.73 (2.01)	-0.60 (0.02)	17.09 (2.10)	**11.96**	-5.18	6.04
stimread	-0.18 (0.03)	-0.78 (1.75)	-0.22 (0.02)	1.81 (1.93)	0.07	0.56	-0.10
workmast	-0.22 (0.02)	7.70 (1.92)	-0.41 (0.02)	14.39 (2.26)	**2.73**	2.74	-1.27
scir	-0.01 (0.02)	17.87 (2.60)	-0.11 (0.02)	22.73 (2.42)	**2.27**	0.53	-0.48
pdir	0.01 (0.02)	15.76 (2.49)	0.17 (0.02)	19.62 (2.09)	**-3.13**	-0.65	0.61
readjoyfreq	0.26 (0.02)	-2.45 (1.49)	-0.08 (0.02)	3.37 (2.38)	1.14	0.46	-1.97
R-Square		0.29 (0.02)		0.25 (0.02)	**15.04 | 23.23%**	**31.06 | 47.98%**	**2.82 | 4.36%**
					The Total Sum of Changes Including the Difference in Constant		**48.92**

Note. Values in bold red represent the significance of the OBD component for each predictor. Coefficients and Interaction terms were not inspected.

Appendix D: Items Used for the Scales

1. joyread

How much do you agree or disagree with these statements about reading?

(Please take into account diverse kinds of reading material, such as books, magazines, newspapers, websites, blogs, emails…)

(Please select one response in each row.)

ST160Q01IA	I read only if I have to.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST160Q02IA	Reading is one of my favorite hobbies.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST160Q03IA	I like talking about books with other people.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST160Q04IA	For me, reading is a waste of time.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST160Q05IA	I read only to get information that I need.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04

Note. Reprinted from PISA 2018 Background questionnaires. 1–83 (p. 35) by OECD (2019f).

2. stimread

In your <test language lessons>, how often does the following occur??

(Please select one response in each row.)

ST152Q05IA	The teacher encourages students to express their opinion about a text.	Never or hardly ever ☐01	In some lessons ☐02	In most lessons ☐03	In all lessons ☐04
ST152Q06IA	The teacher helps students relate the stories they read to their lives	Never or hardly ever ☐01	In some lessons ☐02	In most lessons ☐03	In all lessons ☐04
ST152Q07IA	The teacher shows students how the information in texts builds on what they already know.	Never or hardly ever ☐01	In some lessons ☐02	In most lessons ☐03	In all lessons ☐04
ST152Q08IA	The teacher poses questions that motivate students to participate actively.	Never or hardly ever ☐01	In some lessons ☐02	In most lessons ☐03	In all lessons ☐04

Note. Reprinted from PISA 2018 Background questionnaires. 1–83 (p. 31) by OECD (2019f).

3. workmast

How much do you agree or disagree with these statements about yourself?

(Please select one response in each row.)

ST182Q03HA	I find satisfaction in working as hard as I can.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST182Q04HA	Once I start a task, I persist until it is finished.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST182Q05HA	Part of the enjoyment I get from doing things is when I improve on my past performance.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST182Q06HA	If I am not good at something, I would rather keep struggling to master it than move on to something I may be good at.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04

Note. Reprinted from PISA 2018 Background questionnaires. 1–83 (p. 51) by OECD (2019f).

4. scir
5. pdir

How much do you agree with the following statements?

(Please select one response in each row.)

ST161Q01HA	I am a good reader.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST161Q02HA	I am able to understand difficult texts.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST161Q03HA	I read fluently.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST161Q06HA[4]	I have always had difficulty with reading.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST161Q07HA[5]	I have to read a text several times before completely understanding it.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04
ST161Q08HA[6]	I find it difficult to answer questions about a text.	Strongly disagree ☐01	Disagree ☐02	Agree ☐03	Strongly agree ☐04

Note. Reprinted from PISA 2018 Background questionnaires. 1–83 (p. 40) by OECD (2019f).

4-6. Reverse coded.

6. readjoyfreq

About how much time do you usually spend reading for enjoyment?

(Please take into account diverse kinds of reading material, such as books, magazines, newspapers, websites, blogs, emails…)

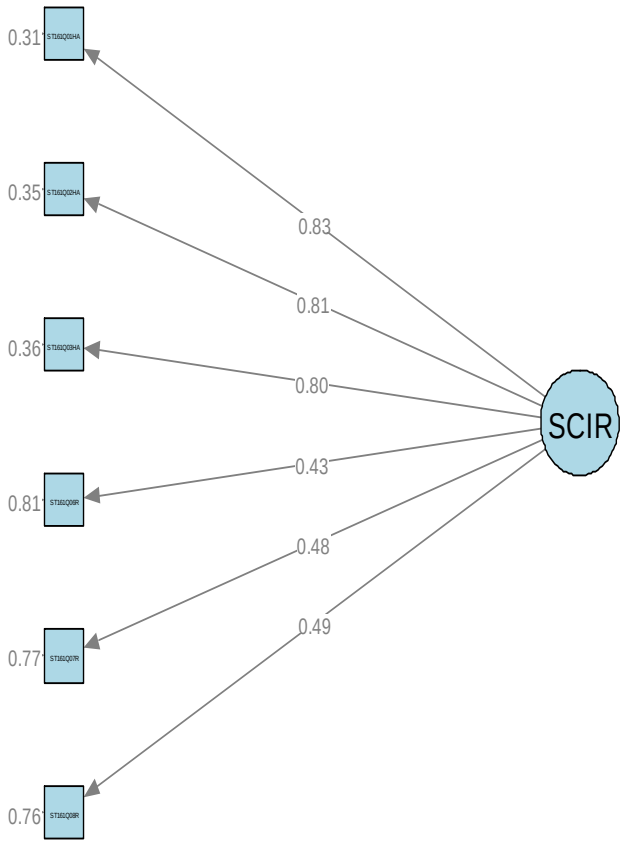
(Please select one response in each row.)

ST175Q01IA	I do not read for enjoyment	☐01
ST175Q01IA	30 minutes or less a day	☐02
ST175Q01IA	More than 30 minutes to less than 60 minutes a day	☐03
ST175Q01IA	1 to 2 hours a day	☐04
ST175Q01IA	More than 2 hours a day	☐05

Note. Reprinted from PISA 2018 Background questionnaires. 1–83 (p. 38) by OECD (2019f).

Figure 2

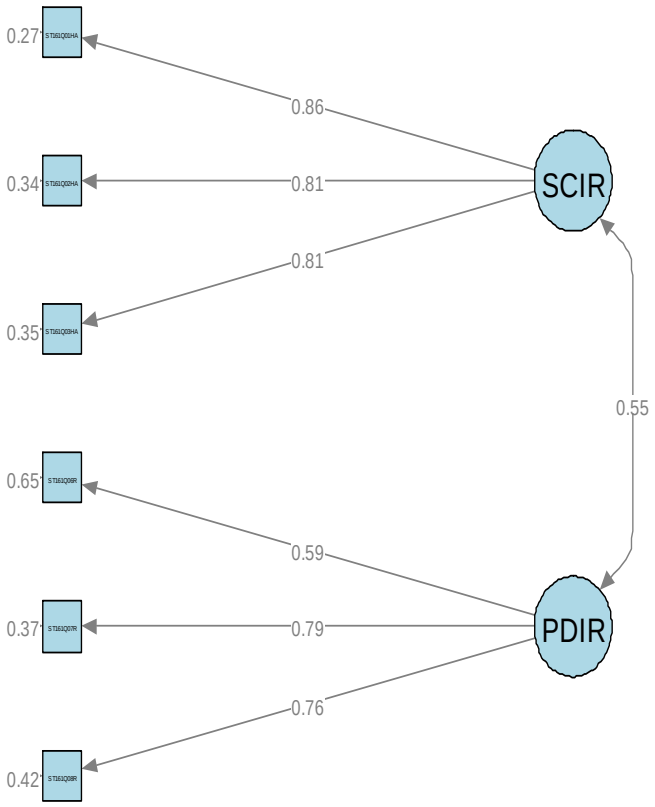
Single-Factor CFA Path Plot



Note. "SCIR" is short for self-concept in reading.

Figure 3

Two-Factor CFA Path Plot



Note. "SCIR" is short for self-concept in reading. "PDIR" is short for perceived difficulty in reading.