

**Equity and excellence in the Kazakhstani education system:
A multilevel analysis of the personal and contextual factors
contributing to students' reading literacy performance on PISA 2018**

Ainur Muratkyzy

Submitted in partial fulfillment of the requirements for the degree of

Master of Science

In

Educational Leadership (School Education)

Nazarbayev University Graduate School of Education

June, 2020

Word count: 20,015 words

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been submitted for the award of any other course or degree at NU or any other educational institution, except where due acknowledgement is made in the thesis. This thesis is the result of my own independent work, except where otherwise stated, and the views expressed here are my own.

Signed:

Date:

AUTHOR AGREEMENT

By signing and submitting this license, I Muratkyzy Ainur (the author or copyright owner) grant to Nazarbayev University (NU) the non-exclusive right to reproduce, convert (as defined below), and/or distribute my submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

I agree that NU may, without changing the content, convert the submission to any medium or format for the purpose of preservation.

I also agree that NU may keep more than one copy of this submission for purposes of security, back-up and preservation.

I confirm that the submission is my original work, and that I have the right to grant the rights contained in this license. I also confirm that my submission does not, to the best of my knowledge, infringe upon anyone's copyright.

If the submission contains material for which I do not hold copyright, I confirm that I have obtained the unrestricted permission of the copyright owner to grant NU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission.

IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN NU, I CONFIRM THAT I HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT.

NU will clearly identify my name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

I hereby accept the terms of the above Author Agreement.

Author's signature:

Date:



Nazarbayev University
Graduate School of Education
www.nu.edu.kz

NU GSE Research Approval Decision Letter

53 Kabanbay Batyr Ave.

010000 Astana,

Republic of Kazakhstan

October 2019

Dear Ainur Muratkyzy

This letter now confirms that your research project entitled: Equity and excellence: A multilevel analysis of Kazakhstani 15-year-olds performance on PISA-2018, has been approved by the Graduate School of Education Ethics Committee of Nazarbayev University. The Committee agreed that your research does not involve the participation of human subjects, and therefore you might proceed with your study.

Yours sincerely,
Daniel Hernández Torrano
Supervisor

On behalf of Elaine Sharplin
Chair of the GSE Research Committee
Professor

Graduate School of Education

Nazarbayev University

Block C3, Room 5006

Office: +7 (7172) 70 9371

Mobile: +7 777 1929961

email: elaine.sharplin@nu.edu.kz

CITI Training Certificate



Completion Date 16-Jul-2019
Expiration Date 15-Jul-2022
Record ID 31864651

This is to certify that:

Ainur Muratkyzy

Has completed the following CITI Program course:

Social & Behavioral Research - Basic/Refresher (Curriculum Group)
Social & Behavioral Research - Basic/Refresher (Course Learner Group)
1 - Basic Course (Stage)

Not valid for renewal of certification through CME. Do not use for TransCelerate mutual recognition (see Completion Report).

Under requirements set by:

Nazarbayev University

CITI
Collaborative Institutional Training Initiative

Verify at www.citiprogram.org/verify/?we090a45f-1cd3-45fa-b9a7-5ce65e3efd17-31864651

Acknowledgements

First of all, I would like to thank myself for being brave enough to accept this challenge of writing a thesis and being hardworking enough to complete it.

I would like to thank my family, especially my mother Almagul for being there for me every day and doing everything possible and impossible to make my life easier and providing me with space to comfortably study as I wished. I love you forever.

I would like to express my sincere gratitude to my supervisor Professor Daniel Hernández-Torrano for supporting me throughout the whole year. He was incredibly helpful and understanding whenever I had trouble balancing my work and study life. I am a big fan of his teaching style and I am ever grateful for having him as my supervisor.

My very special thanks go to Professor Matthew Courtney. From the first email that I have written to him asking for help in conducting my multilevel analysis, he has been nothing but helpful although I took much of his already busy time. Professor Matthew has reviewed and completely rewrote the programming code for conducting my analysis on the R software. I was able to conduct my multilevel analysis only thanks to him.

I have spent amazing two years with my classmates and faculty at NUGSE. They helped me grow both as a person and as a professional. I could not have been here without them.

Last but not least, I would like to thank all of my friends in the “Dasarang” group for providing me with emotional support and helping me relax whenever I needed it.

Abstract

Education plays different roles in governments' agendas in reaching economic prosperity and nations' welfare. In recent years, Kazakhstan took active measures to reform its education system to improve its excellence and equity. In particular, participation in the Programme for International Student Assessment (PISA) in 2009 demonstrated the low level of Kazakhstani students' performance in reading, math, and science literacies. In this context, the relationship between equity and excellence in the education system is an important issue to consider. A little research has been conducted in this field in post-Soviet countries with their unique education system: selective schools for gifted students, bilingual schools, as well as the dominance of STEM subjects in schools. Therefore, the purpose of this quantitative study is to explore the relationship between equity and excellence in the Kazakhstani education system with regards to students' reading literacy based on PISA 2018. More specifically, this study will examine how do students' gender, socio-economic status (SES), schools' gender, SES, language compositions, school type, ownership, and resources affect students' reading literacy. The study used multilevel regression analysis methods at the student and school levels using evidence from PISA 2018. The findings of the study suggest that studying in intellectual schools and schools with the Russian language as a medium of instruction yields an advantage comparable to three years of schooling. Girls were not only found to outperform boys in reading literacy, but also female-dominated schools were found to significantly positively affect reading literacy. Also, students' SES and schools' SES composition were found to affect students' performance, yet school ownership and resources did not demonstrate a significant influence on students' reading literacy in Kazakhstan. This study contributes to understanding the factors behind Kazakhstani students' reading literacy performance in PISA and suggests the possible ways for improving the equity and excellence of the education system.

**Равенство и качество знаний в казахстанской системе образования:
Многоуровневый анализ личностных и контекстуальных факторов,
содействующих повышению уровня
читательской грамотности учащихся в PISA 2018**

Абстракт

Образование играет различную роль в программах правительств по достижению экономического процветания и благосостояния народов. В последние годы Казахстан предпринял активные меры по реформированию своей системы образования с целью повышения ее качества и равенства. В частности, участие в Программе Международной Оценки Студентов (PISA) в 2009 году продемонстрировало низкий уровень успеваемости казахстанских студентов по читательской, математической и научной грамотности. В этом контексте важным вопросом для рассмотрения является взаимосвязь между равенством и качеством знаний в системе образования. Небольшое количество исследований было проведено в этой области в постсоветских странах с их уникальной системой образования: специализированные школы для одаренных детей, двуязычные школы, а также доминирование предметов естественно-математического направления в школах. Поэтому целью данного количественного исследования является изучение взаимосвязи между равенством и качеством знаний в казахстанской системе образования в отношении читательской грамотности учащихся на основе PISA 2018. Более конкретно, в этом исследовании будет изучено, как пол учащихся, социально-экономический статус (СЭС), половой, социально-экономический и языковой состав школы, тип, собственность и ресурсы школы влияют на грамотность учащихся в чтении. В исследовании использовались методы многоуровневого регрессионного анализа на двух уровнях учащихся и школ с использованием данных PISA 2018.

Результаты исследования свидетельствуют о том, что обучение в интеллектуальных школах и школах с русским языком обучения дает преимущество, сравнимое с трехлетним школьным обучением. Оказалось, что девочки не только превосходят мальчиков в грамотности чтения, но и школы, в которых преобладает женский пол, оказывают значительное положительное влияние на грамотность чтения. Кроме того, было обнаружено, что СЭС учащихся и социально-экономический состав школ влияют на успеваемость учащихся, однако собственность школ и их ресурсы не оказывают существенного влияния на уровень читательской грамотности учащихся в Казахстане. Это исследование способствует пониманию факторов, лежащих в основе показателей читательской грамотности казахстанских студентов в PISA, и предлагает возможные пути улучшения равенства и качества знаний в системе образования.

**Қазақстандық білім беру жүйесіндегі білім сапасы мен теңдігі:
PISA 2018 оқу сауаттылығының деңгейін арттыруға жәрдемдесетін
жеке және контекстік факторларды көп деңгейлі анализ жүзінде талдау**

Абстракт

Білім беру халықтың экономикалық өркендеуі мен әл-ауқатын арттыруда үкімет бағдарламаларында әр түрлі рөл атқарады. Соңғы жылдары Қазақстан өзінің білім беру жүйесінің сапасы мен теңдігін арттыру мақсатында реформалау жөнінде белсенді шаралар қабылдады. Атап айтқанда, 2009 жылы Халықаралық Оқушыларды Бағалау Бағдарламасына (PISA) қатысу қазақстандық студенттердің оқу, математикалық және ғылыми сауаттылығы бойынша үлгерімінің төмен деңгейін көрсетті. Бұл тұрғыда білім беру жүйесіндегі теңдік пен білім сапасы арасындағы өзара байланыс маңызды мәселе болып табылады. Бұл салада посткеңестік елдердің бірегей білім беру жүйесінде, атап айтқанда, дарынды балаларға арналған мамандандырылған мектептер, екі тілді мектептер, сондай-ақ мектептерде жаратылыстану-математикалық бағыттағы пәндердің басымдылығы бар елдерде, аздаған зерттеулер жүргізілген болатын. Сондықтан осы сандық зерттеудің мақсаты PISA 2018 негізінде оқушылардың оқу сауаттылығына қатысты қазақстандық білім беру жүйесіндегі білім сапасы мен теңдік арасындағы өзара байланысты зерттеу болып табылады. Нақты айтқанда, бұл зерттеуде оқушылардың жынысы, әлеуметтік-экономикалық мәртебесі (ӘЭМ), мектептің жыныстық, әлеуметтік-экономикалық және тілдік құрамы, мектептің түрі, меншігі және ресурстары оқушылардың оқу сауаттылығына қаншалықты әсер ететіні зерттелетін болады. PISA 2018 деректерін пайдалана отырып, зерттеуде оқушылар мен мектептердің екі деңгейінде көп деңгейлі регрессиялық анализ арқылы талдау әдістері қолданылды. Зерттеу нәтижелері зияткерлік мектептер мен орыс тілінде оқытатын мектептерде оқыту үш жылдық мектеп оқуымен тең артықшылық беретінін дәлелдейді. Қыздар оқу

сауаттылығында ұлдардан жоғары нәтижеге ие болып қана қоймай, сондай-ақ әйелдер жынысы басым мектептер оқу сауаттылығына айтарлықтай оң әсер ететіндігі анықталды. Сонымен қатар, оқушылардың ӘЭМ және мектептердің әлеуметтік-экономикалық құрамы оқушылардың үлгеріміне әсер ететіні анықталды, алайда мектеп меншігі мен олардың ресурстары Қазақстандағы оқушылардың оқу сауаттылығы деңгейіне айтарлықтай әсер етпейді. Бұл зерттеу PISA-да қазақстандық студенттердің оқу сауаттылығы көрсеткіштерінің негізінде жатқан факторларды түсінуге ықпал етеді және білім беру жүйесінде білім сапасы мен теңдігін жақсартудың мүмкін жолдарын ұсынады.

Table of Contents

<i>AUTHOR AGREEMENT</i>	3
<i>NU GSE Research Approval Decision Letter</i>	4
<i>CITI Training Certificate</i>	5
<i>Acknowledgements</i>	6
<i>Abstract</i>	7
<i>Table of Contents</i>	12
<i>List of Tables</i>	14
<i>List of Figures</i>	15
<i>Chapter 1. Introduction</i>	16
1.1 Introduction	16
1.2 Background: Equity and Excellence in Education and PISA	18
1.3 Context: PISA and the Education System in Kazakhstan	20
1.4 Statement of the Problem	22
1.5 Purpose of the Study and Research Questions	24
1.6 Conceptual Framework	25
1.7 Significance of the Study	29
1.8 Outline of the Thesis	30
<i>Chapter 2. Literature Review</i>	31
2.1 Introduction	31
2.2 Previous Studies	31
2.3 Reading Literacy: Equity versus Excellence at the Student-Level	33
2.3.1 Socio-Economic Status	34
2.3.2 Gender	35
2.4 Reading Literacy: Equity versus Excellence at the School-Level	36
2.4.1 School Types	36
2.4.2 Average Class Size	38
2.4.3 School Location and Quality of its Infrastructure	39
2.4.4 Teacher Shortage and Quality	41
2.4.5 Language	42
2.5 Conclusion.....	44
<i>Chapter 3. Methodology</i>	45
3.1 Introduction	45
3.2 Research Design.....	45
3.3 Data	46
3.4 Methods and Procedures	48

3.5 Data Analysis	49
3.5.1 Null model: Student Performance by Schools	50
3.5.2 Model 1: Adding Student-Level Predictors	51
3.5.3 Model 2: Adding School-Level Predictors.....	51
3.5.4 Model 3: Adding All Predictors	52
3.5.5 Model 4: Modelling Significant Predictors.....	52
3.6 Conclusion.....	53
<i>Chapter 4. Findings.....</i>	<i>54</i>
4.1 Introduction	54
4.2 Null Model: Reading Literacy Variation Between Students in Each School	54
4.3 Model 1: The Effects of Student-Level Characteristics on Reading Literacy	56
4.4 Model 2: The Effect of Selected School-Level Variables on Reading Literacy.....	58
4.5 Model 3: The Effect of All School-Level Variables on Reading Literacy	61
4.6 Model 4: Parsimonious Model	65
4.7 Possible Bias in Using Plausible Values	67
4.8 Conclusion.....	67
<i>Chapter 5. Discussion</i>	<i>69</i>
5.1 Introduction	69
5.2 Gender Inequalities	70
5.2.1 Student-Level Gender Inequalities.....	70
5.2.2 School-Level Gender Inequalities.....	72
5.3 Socio-Economic Inequalities.....	73
5.3.1 Student-Level SES Inequalities.....	73
5.3.2 School-Level SES Inequalities.....	74
5.4 School Characteristics and Educational Inequalities	75
5.5 Language Inequalities	77
5.6 Conclusion.....	78
<i>Chapter 6. Conclusion.....</i>	<i>80</i>
6.1 Introduction	80
6.2 Revisiting Research Questions.....	80
6.3 Implications for Theory and Practice	82
6.4 Limitations and Future Research Recommendations.....	83
6.5 Conclusion.....	85
<i>References</i>	<i>86</i>
<i>Appendices</i>	<i>96</i>

List of Tables

Table 1: PISA Results for Kazakhstan Over the Years.....	20
Table 2: Weighted Frequencies and Descriptive Statistics for Independent Categorical Variables.....	47
Table 3: Descriptive Statistics for Independent and Dependent Continuous Variables	47
Table 4: Null Model	55
Table 5: Null Model and Model 1	56
Table 6: Null Model, Model 1, and Model 2	59
Table 7: Null Model, Model 1, Model 2, and Model 3	63
Table 8: Multilevel Models with Main Effects	66

List of Figures

Figure 1: Conceptual Framework..... 26

Chapter 1. Introduction

1.1 Introduction

Education plays different roles in governments' agendas in reaching economic prosperity and nations' welfare. On the one hand, educated youth is often considered as the future skilled workforce driving the economy of the country. On the other hand, education can be seen as a vehicle to overcome poverty and improve the social mobility of citizens. These aspirations represent the ambitions of governments around the world to achieve excellence in education and equity, respectively. More often than not these two goals are seen as contradictory. Some countries focus on funding education for disadvantaged populations, while others let the privileged population benefit from the education system more, decreasing the social mobility. In this regard, educational sociologists debate on the role of education in either maintaining the order in the society or transforming it.

Both equity and excellence are important goals to strive towards. Equitable education system promotes fairness and inclusion that allows all students to receive quality education regardless of their backgrounds and special needs in every classroom. As noble as this goal is, excellence is often seen as a more important goal between the two. Students who excel in schools and become the future quality workforce can drive the economy of the country forward. In the increasingly globalized competitive world, countries often rely on the promotion of excellence, while the issues of equity are disregarded. However, some countries like Canada, Finland, and Japan are successful in promoting both equity and excellence, claiming the tradeoff between them being unnecessary. It is also important to note the difficulty of measuring equity and excellence of an education system. In this regard, the recent developments in conducting international large-scale assessments (ILSA) allow studies of greater depth into the equity and excellence of countries. In particular, Programme for International Student Assessments (PISA) compiles and presents data on a sample population

of 15-year-olds' backgrounds and subject performances from nearly 80 countries worldwide (OECD, 2019c). PISA offers a great opportunity to measure equity and excellence in comparison to other countries.

The Kazakhstani government sees the role of education in improving the country's global competitiveness. The former president Nursultan Nazarbayev shared his vision of education in his annual messages to Kazakhstanis on 5th October of 2018 and 19th March of 2019 and set the path for the education system in country by stating that in order to improve Kazakhstani citizens' welfare, the availability of quality education is a key and that only educated nation could lead the country to become strong economically. The current president Kasym-Zhomart Tokayev continues on the former president's legacy in reforming the education system to improve its quality. The government realizes that equity and country's prosperity go hand-in-hand. Therefore, the extent to which Kazakhstani education system provided equitable opportunities for different subgroups of the population needs immediate attention. These analyses are possible due to the availability of data from ILSA that not only conduct subject tests but also collect information on students' social backgrounds. Kazakhstan is actively participating in ILSA, especially in PISA since 2009. Kazakhstan's growing interest in PISA and other ILSAs is driven by the goal of increasing the global competitiveness of its education system. However, Kazakhstani governments is not only aiming for increasing student excellence, but also for improving equity and inclusion in schools.

Hence, this thesis attempts to establish whether the Kazakhstani education system is providing both equitability and promoting excellence within its education system. The study used quantitative method of research by conducting multilevel analysis of students' individual and schools' contextual factors affecting students' reading literacy.

This chapter aims to provide background information on Kazakhstani education system and how ILSA influenced its current state. The chapter is organized into seven sections. The

first section discusses the equity and excellence debate, and the contribution of PISA in studying this issue. The second section focuses on the case of Kazakhstan in using PISA studies to improve its education system. In the third section, the problem of study is stated. The fourth section outlines the research purpose and research questions. The fifth section states the conceptual framework of the study and lists the variables used in the study. The next section summarizes the significance of the study. The final section outlines the thesis.

1.2 Background: Equity and Excellence in Education and PISA

Among ILSA, PISA played a major role in shaping the current Kazakhstani education system. PISA was first launched in 2000 with 43 country-participants and is repeatedly held every three years. The number of country participants has been growing, with 79 countries taking part in PISA 2018. PISA subject assessments consist of three domains: math literacy, reading literacy, and science literacy. Each test has one subject focus, in the case of PISA 2018 the focus has been on reading literacy, and this is the third time in which reading literacy has been the main focus of PISA.

In OECD terms, reading literacy refers to the “understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society” (OECD, 2019a, p.34). PISA reading literacy tests include tasks that assess students’ ability to engage in higher-order thinking skills by critically analyzing and extracting the key information from considerably large volumes of text. In this era of overload of information on the internet, OECD considers one’s abilities to navigate through them as one of the most important reading skills. Thus, reading literacy differs from the conventional understanding of reading comprehension, which normally refers to students’ ability to read and engage with text. In addition, for PISA 2018, OECD (2019c) reports that most of the 79 country-participants delivered the

test on computers. Thus, for the purposes of the current study, the reference to reading literacy would mean the assessment of reading literacy in a digital environment.

PISA allows not only analyzing students' reading, math, and science literacies; it also provides valuable information on the quality of education in the participating countries. Indeed, PISA data offers invaluable insights into important characteristics of an education system, including issues related to equity and excellence. OECD (2019b) acknowledges the complex nature of equity and defines it through two dimensions: fairness and inclusion. Fairness accounts for the removal of obstacles that can deter students from obtaining educational resources and accessing a quality school environment. In other words, in a fair education system, students have equal access to schools and resources to fulfill their educational needs. Inclusion means that all students, regardless of their backgrounds, have access to high quality education and are able to obtain a basic minimum level of knowledge and skills. However, Roemer and Trannoy denote that equity does not necessarily mean students achieving the same results, but rather every student having an opportunity to realize their potential and acquiring the skills necessary to participate in the society (as cited in OECD, 2019b, p.42). OECD (2014) elaborates that equitable education systems support students in fulfilling their potential without "formally or informally erecting barriers or lowering expectations" (p.17). OECD further add that equity implies SES, gender, ethnicity, family backgrounds not being obstacles in achieving success in education.

The three PISA domains of reading, math, and science literacies are not measured in units and there are no maximum and minimum scores. In fact, PISA scales are expressed in eight levels of proficiency from simplest at level 1c, followed by 1b, 1a, 2, 3, 4, 5 to the most advanced at level 6 (OECD, 2019a, p.43). Students who obtained level 1a and below (i.e., 1b or 1c) in the test are considered to have poor subject knowledge and comprehension. While the

attainment of levels 5 and 6 refers to excellent subject knowledge and comprehension and is considered as excellence (OECD, 2019a). In PISA 2018, approximately 22.6% of all participants obtained level 1a and below, compared to 8.7% of students excelling (OECD, 2019a, p.87). In Kazakhstan's case, 64.2% of participants obtained level 1a and below and 0.4% demonstrated excellence in reading literacy (OECD, 2019a, p.90). As these numbers indicate, there is a huge gap in Kazakhstani students' reading literacy compared to other countries represented in PISA. Hence, this study will provide a novel insight to the literature by exploring the factors that affect Kazakhstani students' excellence in PISA reading literacy.

1.3 Context: PISA and the Education System in Kazakhstan

Kazakhstani 15-year-old students were first tested on PISA in 2009. The country came in 58th place out of 60 country-participants, which was a wake-up call for the government on the poor state of the education system compared to other countries. Since the results were announced, the education system underwent a massive educational reform that introduced fundamental changes in the curriculum, pedagogy, and assessment. In 2012, Kazakhstani students showed a slight improvement in PISA results, especially in mathematics and science literacies. PISA 2015 results were not published due to potential bias in the collected data. In 2018, reading literacy results for Kazakhstan showed no drastic improvements since 2009. Moreover, Kazakhstani students performed worse in all three domains in 2018 compared to 2012 results (see Table 1).

Table 1

PISA Results for Kazakhstan Over the Years

	PISA 2009 Focus on Reading	PISA 2012 Focus on Mathematics	PISA 2018 Focus on Reading
Reading	390	393	387
Mathematics	405	432	423
Science	400	425	397

Source: OECD (2019b, p.307)

Although Kazakhstani government was already taking active steps to reforming the education system by establishing Nazarbayev Intellectual Schools (NIS) in 2008, following the PISA 2009 results, the government initiated a series of changes in the curriculum of schools adopting the best practices of PISA top performers and relying on a University of Cambridge expert team to design a new, updated curriculum (McLaughlin, 2017). One of the most significant changes was the creation of a network of twenty intellectual schools (i.e., Nazarbayev Intellectual Schools; NIS) that would become the ‘agents of change’ and be the first to implement the updated curriculum designed by the University of Cambridge team. The new integrated curriculum aims to develop students’ higher order thinking skills by moving away from the previous content-driven curriculum. Further, the reform involved major changes in the instruction, content, and assessment of students: 1) pedagogical practices were set to become more student-centered, 2) the subject contents were updated in accordance with international practices, 3) formative and summative assessments were introduced to schools, and 4) the government initiated a trilingual policy and prompted STEM teachers to teach in the English language for high school students as stated in the State Programme of Education and Science Development (SPESD) 2016-2019 (2016). These were the ambitious steps policymakers took in order to improve excellence, or PISA scores, and competitiveness of the education system. Equity issues were considered to a lesser extent in these government programmes. Nevertheless, active steps were taken towards developing inclusive education policies. For example, the SPESD outlined the goal to improve school infrastructure in 70% of schools to have conditions for inclusive education by 2019 (p.27).

Equity issues in Kazakhstan are very important due to the diversity and the apparent inequalities in the system. There are two general types of schools in Kazakhstan that differ in their curriculum and financing: mainstream public schools and non-mainstream schools that include lyceums, gymnasiums, schools for children with special

needs, schools for gifted students, and ungraded schools (OECD, 2015). Lyceums, gymnasiums and schools for gifted students benefit from competitive student admission, additional financing from MOE, and more hours of teaching of specific subjects.

Therefore, students with above average abilities in academics prefer to study at non-mainstream schools. During the Soviet era, most of these selective schools admitted students from rural areas and provided them with a boarding school. This way, the Soviet government promoted socialism and attempted to provide equitable opportunities for different subgroups of the population. Nowadays, schools for gifted students have become a competitive ground for urban students despite having boarding schools, as these students can afford private tutoring and can outperform their rural peers in admission tests, thus creating an even bigger gap in their education in the long run (Hernández-Torrano, Tursunbayeva, & Almukhambetova, 2019).

1.4 Statement of the Problem

In many countries, equity and excellence are often seen as being contradictory to one another. However, as Tsai, Smith and Hauser (2017) rightfully point out, “the complementarity between excellence and equity is a morally desirable goal” (p.67). There is a great difference between the average PISA score averaging from some students getting extraordinarily high grades and the rest getting low grades and the students’ obtaining similarly high grades (Schleicher, 2017). Schmidt and Burroughs (2016) also comment on the unnecessary trade-off between equity and excellence. Investing in high performing students with the hopes of higher average performance is not the best available option, since authors provide evidence that countries like Japan, Finland, and Canada were successful in producing “more equity and higher performance” (p.105). According to OECD (2014), “Excellence without equity risks leading to large economic and social disparities; equity at the expense of quality is a meaningless aspiration” (p.15).

The issue of investing in students with ‘more potential to succeed’ could be one of the major inequity issues in Kazakhstan. As mentioned previously, there are several types of schools that benefit from more funds, competitive admission, and advanced curricula in educating the students. In the past years, the 20 NIS schools have received an unproportionally large amounts of funding compared to other schools in the country. Mathematics specialized schools teach math twice as more hours than mainstream schools. The competition for a place in NIS goes up to ten students for one place. The competitive admission becomes a barrier for many students, who cannot afford the private tutoring that goes into preparation for the admission tests (Yakavets, 2014). Despite these barriers-for-entry, according to IAC (2018) statistics, nearly 55% of NIS students are from the socially vulnerable population of the country (p.81). The reason behind such a high number could be twofold since the admission does not account for socially vulnerable population and does not provide any quota for such students. On the one hand, these students could be top performers who “won against the odds”, or on the other hand, the NIS definition of socially vulnerable population might allow for students who are well-off financially. NIS definition of socially vulnerable families includes families from rural areas, single-parent families, large and low-income families, families where parents are either retired or have disabilities, families where children have disabilities, children in care, and orphans. Nevertheless, the system creates a barrier for students who cannot be admitted to these privileged schools and thus are not entitled for the same treatment by the government.

Apart from the equity issues, the Kazakhstani education system as a whole fails to result in student excellence as measured in PISA. Only 0.4% of students, who participated in PISA 2018 reached the levels 5 and 6 in reading literacy, which is considered as student excellence by OECD standards. Even more worryingly, 64.2% of students reached level 1a and below in PISA 2018, which is considered to be the lowest subject knowledge and

comprehension level (OECD, 2019a, p.90). Reading literacy has been the PISA subject domain that Kazakhstani students scored consistently lower compared to math and science literacies. Moreover, Kazakhstan PISA 2018 reading literacy scores were the lowest in all assessment cycles. This result took place despite the country's more than 10 years invested in reforming the education system aiming at improving the functional literacy of students. Therefore, it is also important to study the possible sources of issues in Kazakhstani education system that might have affected the students' reading literacy showing no improvement in 9 years since 2009.

From the above discussion, the type of school students attend in Kazakhstan seems to play a major role in their educational outcomes. Whether students are from opportunity-rich cities like Almaty and Nur-Sultan or from other less developed regions of the country, or whether they attend mainstream or non-mainstream, private or public schools could be a main predictor of their success. Thus, both personal and school contexts are important in studying the social and educational inequalities present in the country. Therefore, this thesis attempts to identify the effect of how much students' personal and schools' backgrounds affect students' PISA reading literacy performance.

1.5 Purpose of the Study and Research Questions

More specifically, the purpose of this quantitative study is to explore the relationship between equity and excellence in Kazakhstani education system with regards to students' reading literacy based on PISA 2018.

Based on the purpose of this study, the main research question of this research is:

- What is the relationship between equity and excellence in Kazakhstani education system with regards to students' reading literacy?

To further explore how personal and social contexts influence students' reading literacy in Kazakhstan, the study aims to answer the following research subquestions:

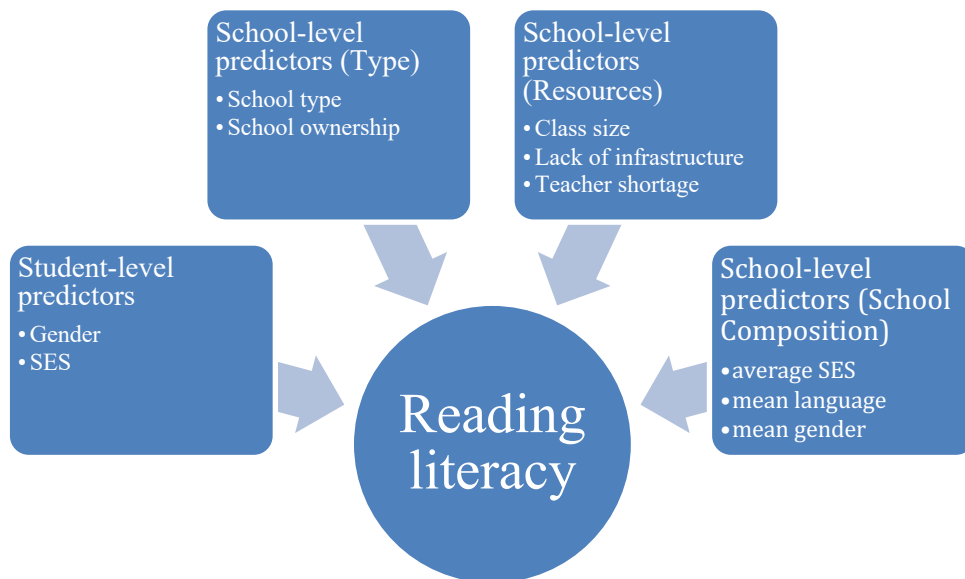
- How do students and schools' composition (i.e., gender, socio-economic status, language) affect students' reading literacy in Kazakhstan?
- How do school type and school ownership affect the students' reading literacy?
- How do school resources (i.e., average class size, school infrastructure, and availability of teachers) influence students' reading literacy?

1.6 Conceptual Framework

The conceptual framework for this study is shown in Figure 1 below. Students' PISA reading literacy scores are expected to be dependent on several student and school-level factors. At the student-level, students' gender and their SES are assumed to be the factors affecting their reading literacy. At the school-level, there are three categories of predictors: type, resources, and school composition. Selective and non-selective school types and private or public schools differing in their ownership affect students reading literacy according to this framework. School resources described by the average classroom size, lack and quality of school infrastructure, and shortage and quality of teaching staff are another category affecting students' achievements. Finally, school composition in terms of students' SES, their gender, and the language of study also impact students' scores in PISA reading. Each of these dependent and independent variables are defined in more detail in the following section.

Dependent variables

The dependent variable in this study is the PISA-2018 scores for reading literacy ('pv1read') since the main domain for the assessment was the reading literacy. Although the official PISA score for the country is estimated using ten plausible values for one domain, the usage of one plausible value will not introduce bias according to the OECD methodology (OECD, 2009). Nevertheless, a further test of nine other plausible values for reading literacy are performed to ensure there is no bias in the analysis.

Figure 1*Conceptual Framework**Independent variables*

As the study aims to evaluate the effects of students' and schools' factors on student performance, predictors at both student and school levels are included in the data analysis. Each of the levels consists of several independent explanatory variables.

Individual-level (level 1). The student-level predictors include gender ('0' for female, '1' for male), and the index of economic, cultural, and social status ('ESCS') score. The ESCS score is a standardized measure presented as an index (thus, missing values in the data were substituted by 0) and is estimated by the OECD (2019a) using the variables of highest occupational status of parents, highest education level of parents in years, home possessions index, which consists of family wealth possessions, cultural possessions, and home educational resources.

There are certainly other explanatory variables for the student-level analysis which might have been useful in predicting the reading literacy of students, however for the current study only two variables are considered. For instance, ethnicity, grade year, and grade repetition are often used for multilevel analysis of equity and excellence relationship, however

in the case of Kazakhstan, most of the students are native and study at 9-10th grades with rare cases of grade repetition, therefore these variables lack an indicative power.

School-level (level 2). The school-level predictors consist of 8 variables: 1) average classroom size ('clsizs'), 2) school type ('sch_int', '0' for non-intellectual, '1' for intellectual), 3) school ownership ('sch_own', '0' for the public, '1' for private), 4) school educational resources, library materials, and school's physical infrastructure index ('edushort'), 5) teacher quality and shortage index ('staffshort'), 6) average school ESCS score or school mean SES ('ESCS.sch'), 7) school mean language ('Language.sch'), and 8) school mean gender ('Gender.sch').

Average classroom size is a variable that is often used as a school characteristic that indicates the quality of school educational resources. In PISA dataset, classroom sizes are averaged for nine ranges, the lowest being 13 for classrooms with 15 students or less and the greatest being 53 for classrooms with 50 students or more.

In countries where schools are segregated into selective and non-selective schools, the school type is found to be a significant predictor of students' academic achievements (Ozdemir, 2016). Similarly, in Kazakhstan, there are selective intellectual schools and non-selective mainstream schools that would be referred to as 'non-intellectual'. Schools are also divided into public and private schools as an important indicator of student and school SES impact on student achievement.

OECD (2019a) introduced two indices for staff and infrastructure quality and shortage for PISA-2018 analysis. The index of educational staff shortage ('staffshort') is a standardized measure and accounts for four items of lack of teachers, lack of teacher assistants, and poorly qualified teachers and teacher assistants. The index of educational material shortage ('edushort') is also a standardized measure and is derived from four items of lack of physical infrastructure, lack of educational materials, and poor quality of physical infrastructure and

educational materials. The ‘edushort’ index represents the state of the school building, heating, cooling, lighting, acoustic systems, IT equipment, library, laboratory, textbook conditions. Positive values for these two indices indicate that a particular school’s principal views the amount and quality of the teaching staff and educational material as an obstacle to a greater degree than the OECD average.

Since Kazakhstan is a plurilingual country, the language used by the students is an important factor to consider. The two official languages in Kazakhstan are the Kazakh and Russian languages. Therefore, students passed the PISA test in only one of the two languages. For the analysis, the ‘language of the assessment’ variable is used as a proxy measure of the language that students predominantly speak. This variable is preferred to the prevalent use of the ‘language at home’ variable in case of Kazakhstan, as language at home can be another language other than Kazakh or Russian, however, students still study in schools either in Kazakh or Russian languages. Even if students speak the Russian language at home, but study in schools in the Kazakh language, their dominant language can be assumed to be Kazakh as they pass the PISA test in the Kazakh language. However, the variable for language cannot be used as a level 1 predictor, since the multilevel analysis require level 1 variables to variate within each school. Monolingual schools, therefore, present the case of no variation of spoken language within schools, which is why the ‘language of assessment’ of students is only be considered as a level 2 predictor.

The three variables of school mean ESCS, school mean language and school mean gender are the variables calculated based on the averages of students within each school. Thus, high school mean ESCS means the predominance of higher SES students within one school. The language variable is coded as ‘0’ for the Kazakh language and ‘1’ for the Russian language. Thus, the school mean language variates between ‘0’ for all-Kazakh language school and ‘1’ for all-Russian language school. Similarly, the gender variable is coded as ‘0’

for male students and ‘1’ for female students, hence the school mean gender variates between ‘0’ for all-male schools and ‘1’ for all-female schools.

1.7 Significance of the Study

The current study is important in establishing the importance of equity in reaching the goal of improving student performance across the country. The benefits of investing in equity have been proven by many studies. According to OECD (2010a) estimates, the potential financial contribution of the OECD countries 15-year-olds reaching at least level 2 in math literacy in PISA could be over USD 200 trillion over their working lives (p.6). OECD (2014) also points out that the benefits of strengthening equity in education include: 1) healthier lifestyles of students; 2) their participation in civil society initiatives; 3) reduction of the cost to society; 4) strong education in early school years leads to easier acquisition and learning of new skills; 5) educating students with disadvantaged backgrounds is also economically efficient in the long term (p.19). Thus, understanding the sources of inequity in the education system and addressing those issues can have a significant impact not only on the education system but also on a country’s welfare.

The current study can most importantly be useful for the government of Kazakhstan and policymakers in understanding and properly interpreting the results of PISA 2018. They will be able to gain insight into the student and school level factors that affect student performance in PISA. In addition, understanding the possible reasons behind Kazakhstani students not showing significant improvements in 9 years in their reading literacy could point at the issues that MOE and other stakeholders need to take into consideration. This could lead to a better planning for PISA 2021 and the overall improvements in the education system of the country. For instance, if school types are found to be a considerable source of inequity, the policymakers might decide to increase the number of schools with advanced curricula to decrease the competition and cater to more students’ needs.

Educational researchers will benefit from the current study since it uses a multilevel analysis method of quantitative research that is rarely used in Kazakhstani ILSA results analyses. This study could be a starting point for other more in-depth quantitative studies based on ILSA comprehensive database for studying Kazakhstani students' performance.

Although there are no direct benefits for specific school communities like school leadership, teachers, and students since this study uses a countrywide data, the results could still be informative and generalizable, and could indicate potential areas for improvement.

1.8 Outline of the Thesis

The following chapter provides a theoretical background of the debate between equity and excellence. Chapter 2 provides a comprehensive literature review of the previous studies on the issues of equity from PISA results and identifying the gaps in these studies. In chapter 3, the methodology of the current study is described in detail. The results and discussion of the results are presented in chapter 4 and chapter 5, respectively. This is followed with a conclusion of the thesis in chapter 6.

Chapter 2. Literature Review

2.1 Introduction

In this chapter, the available literature on student-level and school-level factors that indicate the relationship between equity and students' reading literacy performance on PISA is examined and synthesized. The chapter is divided into three sections. The first section provides a broad overview of the studies conducted on equity and excellence issues in previous years, and the other two sections each focus on different level factors (i.e., student and school levels) affecting equity and excellence in reading achievements of students.

2.2 Previous Studies

The equity and excellence relationship have long interested researchers in education. Education systems around the world aim for providing quality education to their citizens. However, due to numerous individual and contextual circumstances, the education provided to different groups of the population lead to the rise of issues of equity. These issues have been brought to light in educational research journals such as “Equity & Excellence in Education”, which has been publishing studies in this arena since 1963. However, before the Coleman Report in 1966, inequality issues in the US were often assumed to be due to funding differences in the North and the South (Dickinson, 2016). The Coleman Report titled “Equality of Educational Opportunities” was conducted by James Coleman, a sociologist at Johns Hopkins University, by the government’s task to assess the equality of educational opportunities in the US. This was the first turning point in studying the relationship between equity and excellence because Coleman used a self-designed survey to understand what affects students’ learning instead of using government reports on funding. The Report, often deemed as controversial, found that school funding and school resources were not the main influencers of achievement gaps in the American

society, but rather family background and schools' SES composition were found to be more important factors (Dickinson, 2016).

However, the findings of the Report did not resonate well with the political scene of the US at the time and sparked up longstanding debates on educational policies and the relationship between equity and excellence. Still, more than 50 years since its publication, the findings of the Coleman Report seem to hold true. Hanushek (2016) updates on the Coleman Report's findings and states that the effect of family background on students' academic performance has never been disputed since. Many studies expanded on the scope of the research and found similar results in many countries (e.g., Ozdemir, 2016).

The next major shift in studying the relationship between equity and excellence began with OECD launching PISA in 2000 as it compiled comprehensive data on both students' subject performances and questionnaires on students' and schools' backgrounds. Researchers were now able to not only establish the relationship between students' SES and their performance but also take into consideration the contextual factors such as schools' SES, location, leadership, etc. Multilevel regression analysis became one of the major tools in analyzing PISA data since it allowed the grouping of students into 'nests' of schools for studying the effect of schools on student performance.

As mentioned in the previous chapter, PISA tests consist of three subject domains of math, science, and reading literacies, however, studies have been focusing on researching students' performance in one subject domain that has been the focus in a particular year. Thus, at the student level, studies have found that student SES, gender, migrant status, student self-efficacy, and motivation were significant predictors of students' math and science performances in Ireland (Gilleece, Cosgrove, and Sofroniou, 2010), Greece (Karakolidis, Pitsia, & Emvalotis, 2016; Pitsia, Biggart, & Karakolidis, 2017), Turkey (Topcu, Arıkan, & Erbilgin, 2015; Ozdemir, 2016), Singapore (Chen, 2016), Malaysia (Thien, 2016), and also in

cross-country studies (Chiu, 2015). At the school level, school SES, geographical location, teacher quality, and student-teacher ratio were also found to be significant for predicting math and science performances across many countries (Alegre and Ferrer, 2010). However, it is important to note that these findings are not indicative of the relationship between these factors and reading literacy tested at PISA as some studies suggest that different factors affect subject performances to different extents. Tsai, Smith, and Hauser (2017) argue that educational sociologists often wrongly assume that students' social backgrounds similarly affect each subject domain. Hence, further analysis of the literature focus on studies conducted on the analysis of factors affecting reading literacy performance of students.

In summary, the relationship between equity and excellence begun to interest educational researchers since the beginning of the 20th century and has long been studied using different means. The two major turning points in these studies were the controversial Coleman Report published in the 1960s and the launch of PISA in 2000 and subsequent studies using the vast body of data published in these reports. The next sections in this chapter focus on reviewing the most recent findings in studies of equity and excellence in reading literacy primarily based on evidence from PISA studies.

2.3 Reading Literacy: Equity versus Excellence at the Student-Level

In this study, at the student level, only the variables of SES and gender are taken into account as possible predictors of student performance. Certainly, these variables are not exhaustive, and many other factors can affect student performance in PISA. The most often used variables for similar studies are ethnicity or immigrant status, family composition, and grade year (Ozdemir, 2016), however, due to the demographical composition of 15-year-old students in Kazakhstan, these possible explanatory variables are difficult to analyze further and were not considered in the present study.

2.3.1 Socio-Economic Status

Students' social background remains one of the most important predictors of student academic success across many nations. In particular, reading achievements are prone to be impacted by the families' social backgrounds as the performance gap starts as early as early childhood. Studies have found that home literacy resources (Neumann, 2016), early parental involvement in literacy activities (Hemmerechts, Agirdag, & Kavadias, 2017), parents' educational beliefs (Elliott & Bachman, 2018), and slow reading growth at schools struck by severe poverty (Dolean, Melby-Lervag, Tincas, Damsa, & Lervag, 2019) are the factors that differentiated low SES from high SES children's reading literacy achievements.

The social background has an extensive reach towards every aspect of students' lives. In this light, O'Connell (2019) argues that analyzing only students' SES lacks indicative power, which is often assumed when linking students' social class to their academic achievements. Data compiled by ILSA made it possible to measure the effect of students' social background on their assessment results using various items in the questionnaires. For PISA data compilation OECD (2019a) uses an index ('ESCS') for measuring student SES that takes into account the economic, social, and cultural possessions of students. Studies conducted for assessing the PISA reading performance of Albanian (Shera, 2014), Australian (Mahuteau & Mavromaras, 2014), Italian (Agasisti & Longobardi, 2014), Peruvian (Muelle, 2020), Flemish (Rogiers, Van Keer, & Merchie, 2020), and post-socialist Eastern European (Bodovski, Jeon, & Byun, 2017) students, all found an undeniable effect of students SES, with higher SES students getting higher results and lower SES students showing lower levels of reading performance even when controlling for schools' SES. However, there is a lack of studies conducted in post-Soviet countries that could be an example of post-communist economies dealing with social class differences affecting education. Therefore, this study on

Kazakhstani students' SES affecting their reading performance might inform the possible differences between post-Soviet and other countries.

2.3.2 Gender

PISA reading literacy results consistently showed the existence of a significant relationship between gender and reading performance across many countries, with girls outperforming boys (OECD, 2019a). Fredriksson, Holzer, McCluskey-Cavin, and Taube (2009) have found that although the Swedish education system demonstrated a high degree of equity by the lower variance between schools and the lower difference between migrants and native students, the gender difference in students' reading performance in PISA remained high. Furthermore, the authors claim that studies conducted since the 1970s, girls consistently showed better performance than boys in their reading comprehension throughout international comparative studies. The results of these studies seem to indicate that gender differences in reading performance of students are not necessarily the issue of equity and social construction, but rather a phenomenon explained by the nature of reading ability and comprehension (Fredriksson et al., 2009).

Nevertheless, students' social background might deepen the gender differences. Even in earlier years, Cobb-Clark and Moschion (2017) found that among low and middle SES students, girls outperformed boys in reading tests in 3rd grade. Authors attribute this phenomenon to girls being more ready for school at the age of 4 and thus having better literacy skills since kindergarten. Becker and McElvany (2018) discovered similar results during their longitudinal study of students' reading attitudes and behavior in 3rd and 6th grades and found that the gender gap only increased between these periods but less so for higher SES students. This means that although a reading achievement gap occurs from the early ages due to natural gender differences in their development, student SES is an important characteristic that either increases or decreases the gap later.

2.4 Reading Literacy: Equity versus Excellence at the School-Level

Apart from students' social backgrounds, schools' characteristics are also considered to be related to students' achievements. As Alegre and Ferrer (2010) state, institutional differentiation is a mechanism for student segregation that potentially leads to increased inequality in student achievements. Jenkins, Micklewright, and Schnepf (2008) add that this segregation is mainly based on students' social backgrounds rather than schools influencing the segregation. Nevertheless, whether schools enforce segregation or schools are the results of already existing student segregation, the fairness and inclusivity of the education system are at risk. Therefore, it is important to study schools' contextual characteristics and their effect on students' literacy. Especially in the case of Kazakhstan, where schools have major differences in their fees, curricula, teaching staff, and student-teacher ratio, which potentially leads to substantial differences in academic achievements of students.

For this study, the variables of average class size, teacher shortage, lack of infrastructure, school type, school ownership, school mean SES, mean language and mean gender are used. However, other studies use school size and student-teacher ratio as well as the school location either in urban or rural areas (Ozdemir, 2016). From PISA-2018 data, these variables are either unavailable or have a lot of missing or wrong data, which also present problems in their relevance for further statistical analysis. Thus, in this section, only the factors relevant to the current study are reviewed.

2.4.1 School Types

In selective education systems, the issue of inequality between schools and thus inequality of opportunities for students takes a central role. The existence of elite schools either due to their private status or due to the selection of students based on their academic abilities presents a case of widening achievement gaps between different groups of the population. In the education system that segregates students by their gender, race, SES and

performance, equitable and socially transformative functions of the whole system might be compromised. Therefore, it is important to study the effect of school segregation on student performance.

Probably the most open way of segregating students based on their SES is the existence of private schools. The tuition fees for private schools differ from country to country but the mechanism for admission remains fairly similar everywhere: only students with higher SES can afford to study at private schools. However, the achievement gap between students in private and public schools varies across different countries. It is generally accepted that private schools have an advantage over public schools, however, empirical evidence suggests that the academic achievements of students in these two types of schools are similar. Sakellariou (2017) performed analyses on 40 countries that participated in PISA 2012 and found that in only few of countries private schools had an advantage over public schools. Nevertheless, it is still important to note that private schools' location in urban areas and often excellent school resources can affect students' performance in Kazakhstan. Therefore, this study can assist in highlighting the possible issues of students' attending either private or public schools.

Studies show that school segregation is often correlated with the social segregation of the population. However, the attempts to mitigate this social construction yielded conflicting results across different countries. Studies of selective public schools in the US showed that there is a strong effect of attending these elite schools by disadvantaged students, which lead them to attend more selective colleges after graduation (Shi, 2020). A study of selective boarding schools in the UK that admit a high proportion of disadvantaged students has also found that these graduates had higher chances of becoming top performers compared to the control group (Foliano, Green, & Sartarelli, 2019). Whereas in Switzerland, migrant students attending primary schools were found to be subject to discrimination based on their academic abilities and SES, which lead to further reproduction of inequalities (Maag Merki, Buehlmann,

Kamm, Truniger, & Emmerich, 2020). Also, in the UK, Holt and Bowlby (2019) argue that selective schools became elitist and contributed to the hierarchical subjectivities further deepening the social differences of pupils attending them. To which Jerrim and Sims (2019) elaborate that the impact of selective schools in the UK is seen through the divergence of educational aspirations of students between those who attend these schools and those who do not. Based on these findings, it can be concluded that the dominant SES groups within schools can play an important role in predicting its' students' academic achievements.

School competition is another driver of increasing inequalities. Evans (2020) reports that although school districts aim for selective schools being more accessible, school competition drives these schools to admit a certain group of students who will potentially increase the prestige and status of the school. Performance-based selective admittance to schools also serves a similar purpose by admitting only the students who can excel in academics, which further boosts the status of the school. In post-Soviet countries like Kazakhstan, there is a culture of competition for attending selective schools for gifted students. Although these schools are meant for a wider audience and a fair performance-based competition, higher SES students are still more likely to attend these schools due to their ability to hire private tutors and prepare better for admission tests (Yakavets, 2014). Therefore, the inclusion of selective schools in PISA 2018 in Kazakhstan presents an excellent opportunity to study the effect of school type on students' performance based on their social backgrounds and find whether the students belong to a homogeneous group of population.

2.4.2 Average Class Size

There is an ongoing debate on the effect of school resources on student outcomes. A recently published study by Hofflinger and Von Hippel (2020) on low and high SES students' performance in Chile found that family resources had a far greater effect on PISA scores than school resources measured by class size and school competition. Nevertheless, the detrimental

effect of large class sizes had been proven across many disadvantaged populations. In Kenyan secondary schools, teachers and principals cite large class sizes of 40-59 students as one of the major causes of the low quality of teaching and learning (Ndethiu, Masingila, Miheso-O'Connor, Khatete, & Heath, 2017, p.58). South African public secondary school teachers added that large class sizes and lack of resources contribute to the usage of teacher-centered learning methods despite the reforms aimed towards student-centered learning (De Jager, 2017). Especially, this was the case in South African rural multigrade classrooms, where teachers named the class size as the greatest factor that limits their inclusive pedagogy practices (Taole, 2018). The study of Chinese schools in Malaysia has also concluded that among other school-level resources small class sizes contributed to better performance of students on national exams (Kenayathulla, Ling, Razak, & Darusalam, 2019). These findings indicate the existence of the negative effect of large class sizes on student achievement, especially considering the social background effect of students and schools.

In Kazakhstan's case, the issue of class size in different schools and regions of the country might have a strong effect on student performance. There is a general lack of studies on this subject in post-Soviet countries with similar education systems. The operation of multigrade schools in rural areas and specialized selective schools in urban areas might contribute to the considerable achievement gaps among students. Therefore, there is a need to measure the effect of class size on student performance in the context of school segregation.

2.4.3 School Location and Quality of its Infrastructure

In studies of equity and excellence in the education system, the school location and infrastructure issues cannot be disregarded. Great disparities in student achievements have been found among schools based on their location either in disadvantaged neighborhoods in urban areas or their location in rural areas compared to their urban peers. Usually, low-quality infrastructure is the main characteristic of these schools in remote areas (Taole, 2018).

Although the vulnerability of developing or low economy countries is to some extent apparent, the developed countries are not immune to these inequalities in opportunities between regions either. In Parisian middle schools, strong inequalities in test scores between schools in different locations within the city were found (Oberti, & Savina, 2019). France uses the system of school catchment area policy, thus demonstrating how SES of students contributes to the socio-regional and school segregations. Similarly, Shera (2014) found that among Albanian students both gender and SES effects on student performance were associated with the school location among other variables. The same results were reported for Philippine students showing that school location factors exhibited different sensitivity based on students' gender and SES (Trinidad, 2020).

The recent introduction of school reforms is more difficult for schools in rural areas, therefore driving the achievement gap even deeper. Chapman (2020) states that the implications driven by the geographical location of schools are often overlooked by the wider educational stakeholders. He found that teachers in rural areas of Wales were eager to implement the changes introduced in new reforms, however, they faced challenges unique to their location and were unable to adequately respond to the reform. In light of the recent educational reforms in Kazakhstan, the issue of response to the reform of schools in different regions is an important problem to study and could potentially be a new research topic.

Location-wise, Kazakhstan is divided into 3 administrative cities and 14 regions. Despite being centrally managed, each region sets its targets and manages its financial and educational needs locally. Hence, the achievement gap between students from different regions is to be expected. Especially, since the three major cities have the country's greatest number of selective and private schools, possibly indicating a higher SES and higher academic abilities of students residing in these locations. PISA-2018 results show that the mean scores in reading literacy varied between the highest in the city of Nur-Sultan, formerly known as

Astana (428) and the lowest recorded in the Atyrau region (344) (OECD, 2019a, p.75).

According to Woessmann's calculations that 25-30 score points in the PISA scale correspond to one-year-of-schooling (as cited in OECD, 2019a, p.46), the difference between Nur-Sultan city and Atyrau region equals to a staggering three years of schooling. Therefore, it is important to study which of the student and school-level factors are associated with these achievement gaps in Kazakhstan.

2.4.4 Teacher Shortage and Quality

Teacher shortage is undoubtedly a strong obstacle in students' educational achievements. Several studies (Areenpattamannil & Kaur, 2013; Tan & Hairon, 2016; Ndethiu et.al, 2017; Kenayathulla, et.al, 2019) conducted in different countries showed that the number of teachers or the student-teacher ratio were significant factors in predicting student educational outcomes, especially for disadvantaged and immigrant students. Apart from the apparent shortage of educational staff, the qualifications of the existing teachers also play a major role in educating the youth.

Teacher behavior and teaching strategies that promote student learning remain two of the main factors influencing better reading performance. Top PISA performers Singapore and Shanghai based studies showed that teachers' instructional improvements (Huang, Tang, He, & Li, 2019) and teacher-student relationships, classroom disciplinary climate, and teacher stimulation of reading (Ning, Van Damme, Liu, Vanlaar, & Gielen, 2013) contributed to student reading achievements the most. Among Flemish students, teacher involvement and behavior were identified as the greatest factors affecting students' intrinsic motivation to read (De Naegel, Valcke, De Meyer, Warlop, van Braak, & Van Keer, 2014). Therefore, qualified teachers who instill the motivation to read among their students are one of the major factors in improving reading literacy. However, the discourse of equity and excellence requires the

consideration of where these qualified teachers choose to work and how likely it is that higher SES schools hire more qualified teachers.

Qin and Bowen (2019) studied teacher qualification distribution among 32 OECD countries and found that the greatest gaps were found for within-school distributions rather than between-schools. Authors argue that the effect of low SES and high SES schools on the distribution was apparent from the negative impact of income inequalities on the distribution, thus prompting the government action to narrow the gaps by allocating teachers equitably across different SES classrooms. Chiu (2015) also suggests that strategic political intervention to the mechanisms of teacher quality distribution benefits all students despite their SES based on the study of 65 countries participated in PISA. These findings seem to indicate that the natural distribution of teachers is not beneficial in many countries by further deepening the existing social inequalities. Whether this issue similarly manifests itself in Kazakhstan is an important topic to study as there might be quite substantial disparities in income and qualifications between teachers in low SES and high SES schools.

2.4.5 Language

Certainly, the language students speak, and their reading comprehension are closely interlinked. Kazakhstan is an official bilingual country that recognizes the Kazakh language as native and the Russian language as an official language. According to Skutnabb-Kangas (1981) study of bilingualism in the 1980s, although the number of bilingual countries is growing, the greatest issue lies in the differences between linguistic majorities and minorities within these countries. It is important to understand how these language differences manifest themselves in students' reading performance.

Previous studies among multilingual students' reading literacy were mostly conducted in Spain. In Basque Country, where schools provide different models of education based on the bilingualism, bilingual pupils scored the best in the whole country (Sierra, 2008). Contrary

to these findings, Oliden and Lizaso (2014) found that based on PISA-2009 reading results, monolingual Spanish groups of students scored higher than their bilingual counterparts. Nevertheless, Abad and Cabrera-Alvarez (2020) found that if the language at home is either Spanish or one of the co-official languages, then there were no significant differences between these students, but students who spoke other languages at home performed at lower levels. These examples of Spanish and other co-official languages spoken in different regions of Spain demonstrate that in a bilingual society, the differences in reading skills can be mediated and the language spoken at home fails to be an important predictor of students' academic success for linguistic majorities. However, linguistic minorities still suffer from their lack of representation in official assessments such as PISA and is often the cause of lower academic achievements.

There is a general lack of literature on the reading performance of students from bilingual or plurilingual countries. Especially, this is the case for post-Soviet countries, where the Russian language remains one of the common languages spoken at home. Furthermore, the Russian language is predominantly spoken in urban areas rather than in rural areas in Kazakhstan, as well as different regions of Kazakhstan being more "Russianized" than the other ones. Therefore, this study might help to establish whether there is a significant difference between students who attend Kazakh as a medium of instruction schools or Russian as a medium of instruction schools.

In Kazakhstan, the effect of the language of instruction in academic success is illustrated by the results of the Unified National Test (UNT) taken upon graduation from the secondary school in order to enroll in the higher education institution. According to IAC (2018), up to 2016, students passing the test in the Russian language scored consistently higher on average than those passing the test in the Kazakh language (p.105). However, in 2017 and 2018, there is a reverse trend with students passing the test in the Kazakh language scoring higher on

average than their Russian speaking peers. It is important to note that minority languages are not represented in the UNT because school-leavers can only pass the test in either Kazakh or Russian languages, and since 2018 in the English language as well (p.106). Thus, based on UNT results, the language of study does not seem to be an obstacle in meeting students' educational needs for students who predominantly speak the two official languages of the country. However, UNT results can be biased, in which case an international assessment such as PISA can be a better indicator of the educational differences created by the languages of instruction at schools. Therefore, this study is valuable in assessing the effect of the language on students' academic achievements.

2.5 Conclusion

This chapter provided a literature review for each of the individual and contextual characteristics of students and schools that affect students' reading literacy performance within the scope of this study. Although equity and excellence were the topic of research since early 20th century, the launch of ILSA in 2000s changed the way these studies were conducted by compiling a large dataset from different countries. According to these studies, student SES and gender were found to be significant predictors of students' reading abilities. Girls consistently outperformed boys, and students from higher SES families studied better than others. At the school level, studies revealed that the school type, location, resources, and the number and quality of teachers were important factors for enhancing students' motivation and ability to read. However, the effect of language on students' reading is not certain with linguistic majorities performing similarly in bilingual societies, while linguistic minorities performing worse than their peers. Nevertheless, there is a need to study how these factors impact education in a post-Soviet country like Kazakhstan, which presents a unique case of a multilingual country with selective education system. In the next chapter of this thesis, the methodology for measuring those effects are presented and described in detail.

Chapter 3. Methodology

3.1 Introduction

This chapter outlines and justifies the methodological approach used to study the effect of the student and school-level characteristics of Kazakhstani 15-year-olds on their reading literacy performance in PISA-2018. The chapter consists of four sections. First, the research design of the study is outlined. In the second section, the sample for the analysis is described in detail. The third section explains the models used for the multilevel analysis, followed by a concluding section summarizing the chapter.

3.2 Research Design

The purpose of the study justifies the use of quantitative methods to research the relationship between two or more variables (Creswell, 2014). Postpositivist paradigm underlies the ontological and epistemological philosophies of the study, the assumption being that the reality exists but can only be measured and apprehended imperfectly (Guba & Lincoln, 1994). Also, the view of this reality by the investigator is based on the assumption of modified objectivism since the investigator and the subject of investigation are independent entities; hence, the bias from the investigator's part is attempted to be eliminated by following rigorous procedures. According to Guba and Lincoln (1994), the methodology of study under the postpositivist paradigm is based on research questions and hypotheses that are subject to empirical testing but aimed at falsifying hypotheses rather than verifying them.

The study used a descriptive cross-sectional, non-experimental quantitative research design since the study aimed to examine the effect of certain factors on student performance without any experimental manipulations by the investigator. Moreover, the study used secondary data retrieved from PISA database, which was collected at only one point in time.

Studying the effects of student and school factors on student performance independently is impossible since the set of students is within the set of schools. Therefore, the

methodology requires the use of multilevel models to estimate the effects of different level factors: the first level being the student level and the second level being the schools where the students study. In other words, according to Hox, Moerbeek, and Schoot (2018), the data structure is hierarchical or nested and thus requires the use of the multilevel regression analyses to answer the research questions.

There are no ethical issues in conducting this study because the analyses use secondary data from PISA database. Thus, the study does not fall into the category of research involving human participants and does not require ethical considerations.

3.3 Data

The data sample of the research consisted of a population of 15-year-old Kazakhstani students that took part in PISA-2018. A total of 19,507 randomly selected students from 616 schools in 16 regions of the country participated in the assessment (“PISA 2018 Database”, n.d.). The majority of students were students of 9th and 10th grades with some students from 7-8th and 11-12th grades.

Hox et al. (2018) recommend the use of the “50/20 rule” as a rule of thumb to ensure the statistical significance of multilevel models (p.216). According to this rule, the dataset for the analysis should include at least 50 level 2 units (i.e., schools) and each unit should have a minimum of 20 level 1 units (i.e., students per school). While the number of schools in the dataset (616) is more than the recommended 50 schools, a preliminary analysis of the dataset revealed the existence of 181 schools with less than 20 student-participants. Therefore, these 181 schools, which accounted for a total of 1532 students, were removed from the dataset, leaving 435 schools with 17,975 students that would be used for the analysis. Another reduction in the data was performed to account for the existence of single-gender schools in Kazakhstan. The procedures for the multilevel analysis require a variation of variables within each school. However, the variable of gender cannot variate in the case of one-gender schools.

A preliminary analysis showed that there were 3 all-male schools with a total of 125 students. These students were also extracted from the dataset.

After these data reduction procedures, the multilevel analysis for this study used the information on the remaining 432 schools with 17,820 students. The distribution of the remaining sample of students by their gender, language of assessment, and schools is shown in Table 2 below. Descriptive statistics for continuous variables are presented in Table 3.

Table 2

Weighted Frequencies and Descriptive Statistics for Independent Categorical Variables

Categorical variables	N	%	Mean	SD	Min	Max
<i>Gender</i>						
Female	8844	49.6	420.6	84.2	155.8	721.8
Male	8976	50.4	393.3	89.8	163.1	730.5
<i>Language</i>						
Kazakh	10711	60.1	375.3	68.7	163.1	607.1
Russian	7109	39.9	454.4	92.7	155.8	730.5
<i>School type</i>						
Intellectual	2559	14.4	511.6	72.0	297.5	730.5
Non-intellectual	15261	85.6	389.3	77.8	155.8	707.1
<i>School ownership</i>						
Private	1620	9.1	403.6	94.1	211.1	721.8
Public	16200	90.9	407.2	87.5	155.8	730.5

Table 3

Descriptive Statistics for Independent and Dependent Continuous Variables

Continuous variables	Mean	SD	Min	Max
Pv1read	406.7	88.1	155.8	730.5
Clsize	29.9	13.4	13	53
Staffshort	-.487	.981	-1.577	.721
Edushort	-.076	1.114	-1.421	2.960
ESCS. sch	-.298	.449	-1.430	.720
Language.sch	.400	.365	.00	1.00
Gender.sch	1.504	.133	1.07	1.95

3.4 Methods and Procedures

The analysis of factors affecting student performance at both individual and school levels measured in ILSA require advanced quantitative techniques to account for the nested nature of the data. In social sciences, students, apart from being separate entities, are usually also parts of the larger units of classrooms or schools. In such cases, researchers use the hierarchical linear models to account for the school effect on individual student performances by ‘nesting’ students within schools (Raudenbush & Bryk, 2002). Authors denote that hierarchical models have been in use since the 1960’s under different terms of mixed-effects or random-effects models, and multilevel linear or random-coefficient regression models (p.5-6). Nevertheless, the basic principles of the models remain the same.

The current analysis of the sample is carried out using a two-level model, with level-1 being the individual student level that includes the variables of students’ gender and SES, and level 2 being the contextual school level that consists of predictors including school type, school ownership, average class size, lack of infrastructure, teacher shortage, school mean SES, gender, and language of assessment. At each level, multiple explanatory variables help in explaining the variations of the reading literacy performances of Kazakhstani students. The choice of method is justified by the purpose of the study, which aims to measure the effect of students’ factors and schools’ contextual factors on student performance in reading literacy. Thus, the null hypothesis is that student and school-level characteristics of students have no effect on the mean reading achievements of students in Kazakhstan. The null hypothesis is rejected if the student and school-level factors are found to be statistically significant predictors of students’ reading literacy.

The modeling was carried out using the IBM SPSS software and the free online RStudio software (Hox et al., 2018). SPSS software was used for descriptive statistical

analysis and the RStudio software was used for multilevel modeling using various packages built into the programming platform.

3.5 Data Analysis

The data analysis approach for this multilevel analysis uses a variety of statistics. First, the ICC value is used to justify the use of the multilevel regression method. The ICC value indicates the level of clustering within the data. According to Scherbaum and Ferreter (2009), there is no standard plausible value range for ICC that can serve as a means of judging whether there is a substantial clustering present in the data or not. However, based on the ICC value estimates for each country that participated in PIRLS, top performers in the reading domain had lower ICC values within the 0.1-0.3 range (p.261). Hence, for this study, the ICC value above this range is used as an indicator of a substantial between-group clustering present in the data.

Second, the multilevel analyses use regression equations to predict the outcome variable (i.e. reading literacy scores) using different explanatory variables (i.e., independent variables at levels 1 and 2). The regression equation for a two-level model in a combined form is as follows:

$$\text{Reading literacy } (Y_{ij}) = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}W_j + u_{0j} + r_{ij} \quad (1)$$

Equation 1 is adapted from Raudenbush and Bryk (2002) regression equation for a two-level model (p.35-36). Here, $i = 1, \dots, n_j$ are level-1 and $j = 1, \dots, J$ are level-2 units; X_{ij} are level-1 predictor and W_j are level-2 predictor; γ_{00} , γ_{10} , and γ_{01} are the regression coefficients or fixed effects; u_{0j} is the level-2 and r_{ij} is level-1 random effects. The reading literacy outcome (Y_{ij}) is a function of the overall intercept (γ_{00}), the main effects of level-1 and level-2 predictors, and a random error. The intercept, γ_{00} , is the expected reading literacy score of a student with no background characteristics. The slopes, γ_{10} , and γ_{01} , are the expected changes in reading literacy scores of a student with a unit increase in explanatory

variable values. According to Raudenbush and Bryk (2002), the higher the intercept the greater is the school effectiveness or excellence, while the weaker slopes indicate greater equity. The error terms u_{0j} and r_{ij} , are unique effects associated with each school and student.

To obtain efficient estimates of parameters, the full maximum likelihood approach is used. This approach is desirable for this study due to the large sample of data collected from the PISA database, which means that the estimators are approximately normally distributed and are, thus, closer to the true parameter and are “unbiased with minimum variance” (Raudenbush & Bryk, 2002, p.52). The relative fit of the data to the model is tested using the deviance or the $-2\log$ -likelihood, where the likelihood is the value associated with the maximum likelihood estimator under the null hypothesis. According to Raudenbush and Bryk (2002), the large values of the deviance measure can be the evidence of the model being implausible and unfit to describe the data (p.60). Hence, the fit of each of the five models in this study is tested based on their $-2\log$ likelihood, with a smaller value showing a better fit.

The null hypotheses are usually tested at a 5% level of significance; however, the lower p-values at .01 or .001 levels would indicate a highly implausible null hypothesis (Raudenbush & Bryk, 2002, p.71).

The multilevel analyses in this study use five models, each introducing different explanatory variables. The null model or unconditional model explores the student reading literacy scores variation between schools using the ICC value, while Model 1 introduces the level-1 predictors into the model. Model 2 and Model 3 add level-2 predictors to the model. Model 4 is a parsimonious model used for further discussion of the findings. These five models are described in detail in the following sections.

3.5.1 Null model: Student Performance by Schools

The process of the data analysis begins from the preliminary intercept-only model (Model 0) or the null model with no predictors at two levels. The model is run on two

variables: Kazakhstani students' reading literacy performance scores on PISA 2018 ('pv1read') and the schools' unique identification number ('schoolid'). The model estimates the mean reading literacy score or the intercept (γ_{00}) and evaluates the between-school and within-school variation of the dependent variable. These variations are then used to calculate the intra-class correlation (ICC) value, which is the ratio of the between-school variation to the sum of between-school and within-school variations (Scherbaum & Ferreter, 2009). ICC indicates the existence of clustering between schools, or in other words, whether some schools perform significantly better than others. This value informs and justifies the use of multilevel analysis.

Overall, the null model is used to provide evidence of clustering between schools and justifies the use of the multilevel analysis to explain the clustering using predictors at the two levels of students and schools. In the next models, student-level and school-level predictors are included to explain the variations in reading literacy performance within and between schools.

3.5.2 Model 1: Adding Student-Level Predictors

In case the null model shows the presence of clustering between schools, Model 1 is constructed by adding two student-level individual factors to explain the variance of reading literacy within schools at level 1. The level 1 predictors used in Model 1 include students' gender and their individual ESCS scores. The output of Model 1 are the regression coefficients for gender and ESCS indicating the effect of these predictors on students' reading literacy. The null hypothesis for Model 1 is that students' gender and SES have no effect on students' reading literacy.

3.5.3 Model 2: Adding School-Level Predictors

After analyzing level 1 predictors' effect in Model 1, Model 2 further explores the effect of several school-level factors on students' reading literacy performance. In addition to

the level 1 predictors of gender and ESCS, six school-level predictors are included in Model 2: school type ('sch_int'), school ownership ('sch_own'), teacher shortage index ('staffshort'), lack of school infrastructure index ('edushort'), average class size ('clsiz'), and school mean ESCS ('ESCS.sch'). Therefore, a total of eight variables are modeled in Model 2. The outputs are the regression coefficients for each variable. The null hypothesis for Model 2 is that the student and school-level predictors have no effect on students' reading literacy in Kazakhstan.

3.5.4 Model 3: Adding All Predictors

Following the analysis of level 1 and several level 2 predictors in Model 2, Model 3 further models the data by introducing two more school-level factors: school mean language and school mean gender. Thus, Model 3 includes a total of 10 variables: students' gender and ESCS as level 1 predictors, and average class size, school type, school ownership, teacher shortage, lack of school infrastructure, school mean ESCS, school mean language, and school mean gender. The last two variables are added separately in Model 3, rather than in Model 2, to see their effect on other school-level variables from Model 2. If some of the predictors are found to be no longer significant in Model 3, despite being significant predictors in Model 2, this would mean that the school language and gender composition factors were able to explain the variance previously found to be attributed to other predictors. The null hypothesis for Model 3 is again level 1 and level 2 explanatory variables having no effect on students' reading literacy.

3.5.5 Model 4: Modelling Significant Predictors

Although Model 3 is the full model that includes all the variables that were under consideration, Model 4 is constructed as a simpler model that includes only the predictors that were found to be significant for predicting the reading literacy of Kazakhstani students. Model 4 is also called the parsimonious model and is constructed under the assumption of Hox et al.

(2018) that models should generally be kept simpler in order to decrease the deviance, i.e. for the model to fit the dataset better.

3.6 Conclusion

The chapter explained the methodology that is used to measure the effect of students' individual and schools' contextual factors on Kazakhstani students' reading literacy scores in PISA 2018. The multilevel analysis was chosen as an appropriate method for the data analysis since this method allows students to be 'nested' within schools. The total sample of 17,820 students from 432 schools that participated in PISA 2018 is included in the analysis. A series of models were constructed to measure the effects of level 1 (student-level) and level 2 (school-level) predictors. The next chapter presents the findings from conducting the two-level multilevel analysis.

Chapter 4. Findings

4.1 Introduction

The chapter presents the results of the two-level multilevel analysis performed to explore the students' individual and schools' contextual characteristics associated with Kazakhstani 15-year-old students' reading literacy performance in PISA 2018. The main research question asks how schools in Kazakhstan differ in their students' reading literacy performance. Therefore, as described in the previous chapter, a series of five models were run, each including a number of student-level and school-level explanatory variables, testing their significance in predicting reading literacy achievements and measuring those effects. The chapter is organized into seven sections. The first five sections describe the findings from each of the five models separately. The next section evaluates the significance of the findings, and is followed by a concluding section, which summarizes the main findings of the two-level multilevel analysis.

4.2 Null Model: Reading Literacy Variation Between Students in Each School

The null model is run before the analysis to justify the use of the multilevel analysis method. The model allows checking whether there is a substantial clustering between groups of students, or in other words, whether some schools perform significantly better than the others. A significant variation in the reading literacy of students between the schools would indicate a need for further study of students' individual and schools' contextual characteristics that explain the variation in student performance.

The null model is run using only two variables: students reading literacy score ('pv1read') and school's unique identification number ('schoolid'). Table 3 below presents the findings of running the null model for 432 Kazakhstani schools and 17820 students. Here, the intercept is the school mean of the reading literacy score. The regression coefficient for the intercept is 394.0 with a standard error (*SE*) of 2.5. This means that, on average, Kazakhstani

students are expected to achieve a reading literacy score of 394.0 on PISA 2018. Based on PISA 2018 ranking, this score is well below OECD average of 487 in reading literacy (OECD, 2019c, p.17).

Table 4

Null Model

	Model 0 Intercept-only
Intercept	394.0 (2.5)
Units (school/student)	432/17820
-2*log likelihood	-100273.4
	Variation
Between schools	2594.7
Within school	4181.7
ICC	38.3

Table 4 also shows the -2*log-likelihood as well as between-school and within-school variation, which show the relative fit of the model to the data and the residual variation at each level. As explanatory variables are incrementally added to the next models, the gradual decrease in the values of -2*log-likelihood and two-level variations indicate a better fit of the model to the data and the increased variance explained by the predictors.

The intra-class correlation (ICC) is found to be 38.3%. This value is calculated as a ratio of the between-school variability to the total variability. According to Scherbaum and Ferrerter (2009), ICC of 0.25 or 25% for reading performance is an estimate of average ICC for two-level multilevel analysis based on PIRLS studies (p.266). Hence, ICC of 38.3% between Kazakhstani schools informs the presence of substantial clustering and prompts the conduction of the multilevel analysis to explain these variations in reading literacy. Overall, the findings of the null model were used to justify the use of the two-level multilevel analysis. In the next

model, the student-level variables are added to the model and the within-school effects of these variables on reading literacy are measured.

4.3 Model 1: The Effects of Student-Level Characteristics on Reading Literacy

After examining reading literacy variation between students in each school (null model), Model 1 explores the effects of students' individual characteristics on reading literacy within a school. In other words, Model 1 is the first level of the two-level multilevel analysis. The variables included in this model are two student-level explanatory variables: gender and ESCS index indicating students' SES. The 'language' variable could have been considered as a student-level explanatory variable; however, for the multilevel model, the variable should vary within each school, which is not the case for many Kazakhstani schools. Although many schools are mixed language schools, the data shows that out of the 432 schools used in the analysis 206 are monolingual schools, thus not exhibiting any variation within these schools. Therefore, only gender and ESCS effects are measured in this model.

Table 5

Null Model and Model 1

	Model 0 Intercept-only	Model 1 Within-school effects model
Intercept	394.0 (2.5)	429.6 (2.8)
Gender (Male)		-22.0 (1.0)**
ESCS		6.1 (0.7)**
Units (school/student)	432/17820	432/17820
-2*log-likelihood	-100273.4	-99989.7
	Variation	
Between schools	2594.7	2393.0
Within school	4181.7	4056.5

Note. ** $p < .01$; -2*log-likelihood indicates a relative fit of the model to the data.

Table 5 shows that both predictors are significant at a 1% level. After introducing these variables, the intercept increased to 429.6 (SE = 2.8). Thus, a randomly selected Kazakhstani 15-year-old student is expected to achieve a score of 429.6 in their reading literacy. Now, using the Model 1, students' gender and SES can be used to predict students' achievements.

First, the regression coefficient for the gender variable is -22. The negative coefficient for gender means that male students (coded as '0') perform significantly worse than female students (coded as '1'). The difference between scores equals 22 points in the PISA test, meaning that if the randomly chosen student is a boy, 22 points are subtracted from the intercept to achieve their expected PISA score in reading literacy. In other words, female students are expected to achieve a score of 429.6 (the intercept), while male students are expected to achieve 22 points less or 407.6 points. Usually, 25 points in PISA are associated with a one-year of schooling difference (OECD, 2019a), which allows the interpretation of the gender achievement gap between students as nearly a one-year of schooling difference.

Second, the ESCS is found to be positively correlated with the students' reading literacy achievements, which is often the case in many countries. This means that higher SES students are expected to achieve higher scores in PISA. However, compared to gender, the effect is much smaller for SES. While the gender achievement gap could be compared to 22 points in PISA, higher SES gives an advantage of only 6.1 points on average.

Looking at the relative fit of the model to the data from Table 5, it is clear that the $-2 \times \log$ -likelihood decreased compared to the null model, meaning that Model 1 fits the data better than the null model with no predictors. The between-school and within-school variations also decreased, which shows that the two student-level variables explain the previous residual variation. Overall, the findings for Model 1 indicate that gender and ESCS are significant predictors of Kazakhstani students' reading literacy. The next model further introduces the school-level predictors into the model.

4.4 Model 2: The Effect of Selected School-Level Variables on Reading Literacy

After examining the significance of students' individual level characteristics on their reading literacy in Model 1, Model 2 explores the effects of school level contextual factors on reading literacy differences between schools. On top of gender and ESCS factors, six school-level explanatory variables are included in Model 2, namely, average class size ('clsize'), school type ('sch_int'): intellectual and non-intellectual schools, school ownership ('sch_own'): private and public schools, school average ESCS score ('ESCS.sch'), teacher shortage index ('staffshort'), and lack of infrastructure index ('edushort'). Thus, Model 2 considers both individual level predictors and school level predictors in modeling the data. Table 6 presents the findings for Model 2 in comparison to the null model and Model 1. Overall, the findings indicate that Model 2 is a better fit for the data than Model 1 since the $-2 \times \log\text{likelihood}$ saw a reduction and residual variance at the school-level reduced nearly twice. This means that the school-level predictors were significant in explaining the between-school variance of the reading literacy achievements.

Individual-level predictors (level 1). Firstly, the effects of gender and ESCS at the student-level are found to be significant at a 1% level, but with a slight decrease in their effects, down to -21.8 for gender and 5.2 for ESCS. The intercept for Model 2 increased to 482.1. Thus, according to Model 2, a randomly selected Kazakhstani student is expected to achieve a score of 482.1 points in their reading literacy. To get an expected score for male students, 21.8 points should be subtracted from the intercept. Thus, a randomly selected female student would achieve a score equal to the intercept (482.1), while a male student would achieve 21.8 points less or 460.3 points on average. For higher SES students, a further 5.2 points should be added to calculate the expected score. Overall, after adjusting the model for school-level variables, the individual level 1 effects remain similar to Model 1.

Table 6*Null Model, Model 1, and Model 2*

	Model 0 Intercept-only	Model 1 Within-school effects model	Model 2 Between-school effects model
Intercept	394.0 (2.5)	429.6 (2.8)	482.1 (9.2)
<i>Student-level</i>			
Gender (Male)		-22.0 (1.0)**	-21.8 (1.0)**
ESCS		6.1 (0.7)**	5.2 (0.7)**
<i>School-level</i>			
Clsize			-0.6 (0.1)*
Staffshort			0.5 (2.0)
Edushort			0.2 (2.0)
Sch_int (Intellectual)			68.0 (9.6)**
Sch_own (Public)			-14.8 (6.6)
ESCS.sch			57.7 (4.9)**
Units (school/student)	432/17820	432/17820	432/17820
-2*log-likelihood	-100273.4	-99989.7	-99848.6
Variation			
Between schools	2594.7	2393.0	1267.7
Within school	4181.7	4056.5	4057.0

Note. * $p < .05$; ** $p < .01$; -2*log-likelihood indicates a relative fit of the model to the data.

School-level predictors (level 2). Level 2 factors affecting the reading literacy of Kazakhstani students for Model 2 include the average class size, school type, school ownership, teacher shortage, lack of infrastructure, and mean school SES.

The average class size ('clsize') is found to be a significant predictor of reading literacy achievements at a 5% level, the regression coefficient being -0.6. The negative sign of the coefficient means that the smaller the class size the higher are the scores are for reading

literacy. Compared to the effect of level 1 predictors gender and ESCS, the effect of the class size is rather small.

The school type ('sch_int') is found to be a significant predictor of student achievement at a 1% level, showing the effect of 68.0 points advantage of intellectual schools over non-intellectual schools. Since intellectual schools are selective schools, their positive effect on academic achievements is expected. The difference in reading literacy scores between intellectual and non-intellectual schools is approximately three years of schooling, which indicates a considerable gap between the two types of schools.

Surprisingly, both indices for teacher shortage ('staffshort') and lack of infrastructure ('edushort') show a small positive correlation and are not statistically significant predictors of Kazakhstani students' reading literacy. These results potentially indicate that if one school has a shortage of teachers and both a lack and low quality of school infrastructure, the students would perform at a similar level to those who attend schools with no shortage of teachers and good school infrastructure. However, careful consideration in interpreting these results is needed since the school type variable has a great effect on student achievements. Intellectual schools are usually the ones with low staff shortages and excellent school infrastructure. Thus, further analysis are needed to determine whether the school type variable can explain the majority of the difference between schools' staff and infrastructure effects on the students' reading literacy.

The mean school SES or ESCS index is the average of ESCS scores of students within each school. Table 5 shows that this predictor has also accounted for one of the biggest statistically significant effects on the students' reading literacy. The regression coefficient of 57.7 indicates that the prevalence of higher SES students within a school could lead to a 57.7-point gain in reading literacy scores on average.

Another outcome of the multilevel analysis is the school ownership on the students' reading literacy being statistically not significant. The regression coefficient is equal to -14.8, indicating the existence of the negative effect of studying at public schools rather than private schools. Since the predictor is not statistically significant, it is safe to assume that, on average, students see no gain from attending private schools according to the analysis. However, once again, careful interpretation of these findings is needed, as the analysis showed that higher mean SES schools perform better than those with a lower mean SES schools. Therefore, the difference between public and private schools might be to some extent attributable to the effect of the variable of the mean school SES ('ESCS.sch').

Overall, findings derived from Model 2 indicate that the student level predictors remain significant after adjusting for school-level variables, and that average class size, school type, and mean school SES are significant predictors of Kazakhstani students' reading literacy. More specifically, the findings show that students in smaller classes, as well as those attending intellectual schools, and schools with higher mean SES, obtain higher scores than students in larger classes, non-intellectual schools, and schools with lower mean SES. However, students' reading literacy scores do not seem to be affected by other school factors such as the ownership of the school (private vs public), teacher shortage, and lack of infrastructure. In the next model, another two school-level predictors of school mean for language and gender are added. They were not included in Model 2 to see their effect on other school-level variables, as language and gender are important characteristics affecting students' reading literacy.

4.5 Model 3: The Effect of All School-Level Variables on Reading Literacy

After examining the effect of gender, individual SES, average class size, school type, school ownership, teacher shortage, lack of school infrastructure, and school mean SES on students' reading literacy in Model 2, Model 3 adds two more school-level predictors: mean

school language and mean school gender. The meaning of mean school language and gender might be difficult to comprehend in a real-life context, but they can be referred to as the prevalence of one language over another and the prevalence of one gender over another in a given school. Thus, Model 3, or the full model, includes all the variables under consideration in this study, two student-level predictors and eight school-level predictors.

Table 7 shows that the full model is a better fit for the data compared to the previous models indicated by the reduction in the $-2 \times \log$ -likelihood and nearly three-times the reduction in the residual between-school variation.

Individual-level (level 1). The within-school variation is at a similar level as in previous models, which is expected since no additional student-level variables were included in Model 2 and Model 3. Therefore, it is not surprising that the regression coefficients for gender and ESCS remained stable. The intercept for Model 3 is 453.2, seeing a reduction of almost 30 points from Model 2. Given that individual-level effects remained the same, smaller intercept results in lower expected scores for students' reading literacy at level 1 compared to Model 2 expectations.

School-level (level 2). The number of school-level predictors increased for Model 3 and include the average class size, school type, school ownership, teacher shortage, lack of infrastructure, and mean school SES, language, and gender.

The effect of studying in an intellectual school became even larger in this model. The regression coefficient is 77.5, significant at a 1% level. Considering the intercept in Model 3 is smaller than in Model 2, in terms of PISA score expectations, studying in an intellectual school yields a lower expected value in Model 3 than in Model 2. This means that a randomly selected female student studying at intellectual school would be expected to achieve 530.7 points according to Model 3, while according to Model 2 she should have achieved 550.1 points. Thus, the numerical increase in the effect of the school type is an expected outcome of

adding new predictors to the model. Nevertheless, the difference of 77.5 points between intellectual and non-intellectual schools is substantial as it is equivalent to three years of schooling.

Table 7*Null Model, Model 1, Model 2, and Model 3*

	Model 0 Intercept-only	Model 1 Within-school effects model	Model 2 Between-school effects model	Model 3 Full model
Intercept	394.0 (2.5)	429.6 (2.8)	482.1 (9.2)	453.2 (14.5)
<i>Student-level</i>				
Gender (Male)		-22.0 (1.0)**	-21.8 (1.0)**	-21.8 (1.0)**
ESCS		6.1 (0.7)**	5.2 (0.7)**	5.2 (0.7)**
<i>School-level</i>				
Clsize			-0.6 (0.1)*	0.05 (0.1)
Staffshort			0.5 (2.0)	-1.0 (1.3)
Edushort			0.2 (2.0)	0.2 (1.3)
Sch_int (Intellectual)			68.0 (9.6)**	77.5 (6.1)**
Sch_own (Public)			-14.8 (6.6)	-3.9 (4.4)
ESCS.sch			57.7 (4.9)**	46.0 (3.4)**
Language.sch				76.6 (3.2)**
Gender.sch				-25.0 (8.9)*
Units (school/student)	432/17820	432/17820	432/17820	432/17820
-2*log-likelihood	-100273.4	-99989.7	-99848.6	-99664.9
Variation				
Between schools	2594.7	2393.0	1267.7	484.9
Within school	4181.7	4056.5	4057.0	4057.4

Note. * $p < .05$; ** $p < .01$; -2*log-likelihood indicates a relative fit of the model to the data.

The prevalence of Russian language learners ('Language.sch') in a school had a significant effect on students' reading literacy scores. As shown in Table 6, the regression

coefficient is 76.6 and is statistically significant at a 1% level. This value shows the positive effect of studying in a predominantly Russian language speaking school on reading literacy performance. Interestingly, the effect of the language is similar to the effect of attending an intellectual school, 76.6 and 77.5 respectively. In other words, attending a non-intellectual school and a school with predominantly Kazakh language speaking students are similar for Kazakhstani schools. The possible reasons behind Russian language speakers' advantage being comparable to studying in a selective school are further discussed in the next chapter.

On top of gender differences at the individual student-level, the analysis shows that the gender gap increases at the school level. The predominance of male students in a given school has a statistically significant (at 5% level) negative effect on students' reading literacy achievements in PISA 2018. This means that a female student attending a predominantly female school is expected to achieve a score equal to the intercept (453.2) and a male student attending a predominantly male school is expected to achieve 21.8 points less, due to within school gender differences, and another 25 points for between school gender differences in reading literacy performance. Thus, gender alone can accumulate almost a two-years of schooling difference in reading literacy in Kazakhstan.

The significant effect of the average class size has disappeared in Model 3 after adjusting the model for school mean language and gender. The regression coefficient of 0.05 indicates a small and not statistically significant positive effect. Class size does no longer predict students' reading achievements after accounting for the language and gender variations between schools. This indicates that these two new variables already account for the average class size. The correlation between them is not evident at the first glance and needs a further study into the effect of the predominance of one language and one gender within schools to understand how the average class size effect can be explained by them.

The two indices of ‘staffshort’ and ‘edushort’ remain to be not statistically significant. Table 6 shows that the effect of ‘staffshort’ became negative compared to Model 2, which indicates that the shortage of teachers resulted in better student performance. However, since the value is not statistically significant, this negative effect cannot be generalized for the whole population.

The effect of public schools became smaller in Model 3 after introducing the school mean language and gender predictors. Thus, the regression coefficient equals -3.9, which is still not statistically significant. However, the effect of school mean SES is statistically significant, accounting for 46.0 points in score difference. In other words, the predominance of higher SES students in a school gives the advantage of 46 points or nearly two-years of schooling in the reading literacy scores compared to studying in a school with predominantly lower SES students. Therefore, despite the model showing no significant differences between private and public schools, the predominance of a certain group of students based on their SES is still an important predictor of student academic success.

Overall, the Model 3 findings show that at the individual level gender and SES are significant predictors of students’ reading literacy and school type, school mean SES, language, and gender are significant predictors at the school level. Based on these results, a parsimonious model is considered in the next section.

4.6 Model 4: Parsimonious Model

The last model analyzed in this study is the parsimonious model, which includes the variables that were found to be statistically significant predictors in the previous models. Therefore, this model includes only six variables: two student-level and four school-level predictors: gender, ESCS, school type, school mean ESCS, school mean language, and school means gender. The model fits the data similarly as Model 3, as the $-2 \times \log$ likelihood and

within-school and between-school variations remain at similar levels. The intercept slightly decreased to 451.6 compared to Model 3.

Table 8

Multilevel Models with Main Effects.

	Model 0	Model 1	Model 2	Model 3	Model 4
	Intercept-only	Within-school effects model	Between-school effects model	Full model	Parsimonious model
Intercept	394.0 (2.5)	429.6 (2.8)	482.1 (9.2)	453.2 (14.5)	451.6 (13.1)
<i>Student-level</i>					
Gender (Male)		-22.0 (1.0)**	-21.8 (1.0)**	-21.8 (1.0)**	-21.8 (1.0)**
ESCS		6.1 (0.7)**	5.2 (0.7)**	5.2 (0.7)**	5.2 (0.7)**
<i>School-level</i>					
Clsize			-0.6 (0.1)*	0.05 (0.1)	
Staffshort			0.5 (2.0)	-1.0 (1.3)	
Edushort			0.2 (2.0)	0.2 (1.3)	
Sch_int (Intellectual)			68.0 (9.6)**	77.5 (6.1)**	76.4 (5.8)**
Sch_own (Public)			-14.8 (6.6)	-3.9 (4.4)	
ESCS.sch			57.7 (4.9)**	46.0 (3.4)**	46.2 (3.4)**
Language.sch				76.6 (3.2)**	76.3 (3.1)**
Gender.sch				-25.0 (8.9)*	-25.3 (8.8)*
Units (school/student)	432/17820	432/17820	432/17820	432/17820	432/17820
-2*log-likelihood	-100273.4	-99989.7	-99848.6	-99664.9	-99668.9
Variation					
Between schools	2594.7	2393.0	1267.7	484.9	481.5
Within school	4181.7	4056.5	4057.0	4057.4	4057.5

Note. *p < .05; **p < .01; -2*log-likelihood indicates a relative fit of the model to the data.

Table 8 summarizes the findings from all the models analyzed in this study. The gender and ESCS effects also remain the same as in previous models. The effects of intellectual schools and mean school language became almost equal, 76.4 and 76.3,

respectively. The effects of mean school ESCS and mean school gender slightly increased in decimals to 46.2 and -25.3, respectively. Thus, it can be concluded that Model 3 and Model 4 yield similar results, meaning that gender, individual SES, school type, school mean SES, school mean language, and school mean gender are significant predictors of Kazakhstani students' reading literacy performance in PISA.

The Model 4 is the final model used to analyze the effect of different factors at student and school levels on students' reading literacy on PISA 2018. The next section describes an additional set of analyses aimed at checking whether the plausible values used for the multilevel analyses are unbiased, since OECD provides ten plausible values for each subject domain results of students.

4.7 Possible Bias in Using Plausible Values

This study employed first of the ten plausible values for reading literacy performance of Kazakhstani students. Further analyses were carried out in order to ensure that using other plausible values would have yielded similar results. The code on R was duplicated and 'pv1read' values were consecutively replaced by the other nine plausible values. Each time the final table of regression coefficients and the variables that were statistically significant remained the same. Therefore, it is possible to conclude that 'pv1read' is an unbiased measure of students' reading literacy scores on PISA.

4.8 Conclusion

In the chapter, I presented the findings from the two-level multilevel analysis performed on the dataset consisting of Kazakhstani 15-year-old students who passed the PISA test in 2018. A series of five models were constructed, each introducing additional explanatory variables. At the student-level, two variables were found to have a statistically significant effect on reading literacy scores: gender and SES. Female students were found to perform significantly better than boys, the score difference being equivalent to one-year-of-schooling.

Also, SES acted as a statistically significant predictor of the dependent variable. More specifically, higher SES students seem to perform slightly better than lower SES students. At the school level, the two biggest effects are attributed to the school type and mean school language. Students attending intellectual schools and students studying at schools with predominantly Russian language speakers both had an advantage equivalent to three-years of schooling over the students attending non-intellectual schools and those studying at schools with predominantly Kazakh language speakers. Mean school SES also had a positive effect on students' reading literacy scores, with students attending schools with predominantly high SES students performing better than those attending schools with predominantly low SES students, accumulating a gap of nearly two years of schooling. Finally, mean school gender was also found to be a significant predictor of students' reading achievements, with students attending schools with predominantly female populations performing better than students attending schools with the predominantly male populations by a score equivalent to a one-year of schooling. In the next chapter, these findings are further discussed in more detail.

Chapter 5. Discussion

5.1 Introduction

The previous chapter described the findings from the two-level multilevel analysis on students' and schools' data for 17,820 Kazakhstani 15-year-olds from PISA 2018 with regards to the students' reading literacy. The study aimed at evaluating the extent to which Kazakhstani students' personal and schools' contextual factors affected the outcomes of the PISA 2018 reading literacy results, or in other words, assessing the relationship between equity and excellence in Kazakhstani education system. The data analysis focused on measuring the effects of student and school level factors on student reading literacy. Student level factors included students' gender and SES, while school level factors included school mean gender, SES, language, school type, school ownership, average class size, school infrastructure, and teacher shortage.

In general, the findings of this study suggest that although student-level factors are important in estimating Kazakhstani students' reading literacy, school-level factors of school mean language and school type exhibit a far greater effect on students' reading literacy. In this chapter, findings are discussed in the light of the previous literature in broad categories of gender inequalities, SES inequalities, school characteristics and educational inequalities, and language inequalities. The implications of these findings are intended to be used in explaining the equity and excellence issues underlying PISA 2018 scores obtained by Kazakhstani students.

The chapter is organized into five sections. In the first section, student and school mean SES effects on reading literacy are discussed in detail. In the second section, the gender issues in reading literacy are discussed on individual and school levels. The third section focuses on school type, school ownership, and school resource characteristics and the educational inequalities they affect. The fourth section gives a detailed account of the

language issues affecting Kazakhstani students' reading literacy. The chapter concludes with a conclusion section summarizing the main arguments.

5.2 Gender Inequalities

Gender inequalities in student achievements is a commonly researched topic in the studies of equity and excellence. PISA scores have often brought light to these gender achievement gaps in students' reading, math, and science literacies. In the Kazakhstani students' case, gender was found to be a significant predictor of students' reading literacy in PISA 2018 both at the individual and school levels. In this section, the possible explanations and interpretations of these findings are offered.

5.2.1 Student-Level Gender Inequalities

The parsimonious model of the multilevel analysis showed the difference of approximately 22 points between girls and boys in PISA reading literacy in Kazakhstan. Based on Woessmann's interpretation, 25 points in PISA are equivalent to one-year-of-schooling (as cited in OECD, 2019a, p.46). Thus, a 22-point difference can be compared to almost one year of schooling, which indicates that boys in Kazakhstan are approximately one school year behind girls in their reading literacy. This finding refers to a countrywide issue of male students underperforming in reading literacy compared to female students. It might indicate a greater issue in the education of students with possible implications on future educational achievements of these students. There are a number of possible reasons that may explain these gender differences found in Kazakhstani students' reading literacy. First, there is strong empirical evidence that suggests the existence of gender differences in PISA reading literacy across many countries. Thus, the findings of this study are consistent with other studies on PISA reading literacy, where girls were found to significantly outperform boys (OECD, 2019c). Despite these consistent findings in many countries, the possibility of the social construction of gender differences cannot be ruled out, especially since the effect size

differs from country to country. As suggested by Becker and McElvany's (2018) study, these might be attributed to students' SES differences where the gender gap in reading achievements is smaller for higher SES groups of students. Kazakhstan's contextual factors might also be important in explaining the gender gap because, since Soviet times, the country's education system prioritized STEM subjects over the others and male students were expected to succeed in these areas of study leaving them to regard non-STEM subjects as unimportant.

In order to enhance the understanding of gender issues among Kazakhstani students, similar multilevel analyses were performed on Kazakhstani students' math and science performance on PISA 2018 (see Appendix A and Appendix B). Interestingly, reading literacy results were found to be the most sensitive to the gender predictor compared to math and science literacies in Kazakhstan. Boys were found to perform better than girls in math literacy with a statistically significant difference of 6 points in PISA scores. However, girls were still found to outperform boys in science literacy by 4 points. These regression coefficients are much smaller compared to the 22-point difference in reading literacy. These gender differences in math and science are to some extent consistent with the literature where being a male student is a significant predictor of math performance, while the gender preference varies for science performance for different countries (OECD, 2019c). However, the nearly 4-5 times difference between the effect of gender on reading over math and science is an important issue to consider. These findings might suggest that 1) the excessive attention given to STEM subjects in recent years resulted in boys and girls coming closer to closing the gap in math and science literacies compared to reading literacy (Van Langen, Bosker, & Dekkers, 2006); 2) girls are naturally more inclined to perform better in reading from early childhood (Cobb-Clark & Moschion, 2017); 3) a gender gap seems to manifest itself differently depending on student SES (Becker & McElvany, 2018). Apart from these individual level differences

between students with different gender, there are also school level gender gaps present in Kazakhstan, which are discussed further.

5.2.2 School-Level Gender Inequalities

The female student dominance in schools was also found to be a significant predictor of reading literacy of Kazakhstani students. The gap amounted to nearly two years of schooling. This finding refers to the existence of greater gender achievement gaps based on school's gender composition. It is important because according to these findings, male students are at a far greater disadvantage in terms of their reading literacy achievements than previously anticipated by the individual level gender differences. School gender composition is one of the relatively less researched areas in the studies of equity and excellence. Therefore, the possible reasons behind this finding are difficult to explain and interpret. Nevertheless, to offer more insight into the issue, a similar analysis for math and science literacies were conducted. In both cases of math and science, the prevalence of female students in schools had a positive but a statistically non-significant effect on math and science performances. In other words, the prevalence of male students in schools was not advantageous for both STEM and non-STEM subject achievements. According to a cross-national study by Van Langen et al. (2006), two plausible reasons may explain these findings. First, girls are increasingly improving their relative achievements against boys in secondary education and show an increasing participation in tertiary STEM education. Second, this gap may also be associated with the finding in Van Langen et al. (2006) finding that integrated education systems favor girls' achievements compared to differentiated education systems. Thus, Kazakhstan's gradual changes to a differentiated education system might lead to a smaller effect of gender on students' reading literacy.

5.3 Socio-Economic Inequalities

The discussion of equity and excellence issues is impossible without consideration of the students' socio-economic status, or, based on the OECD conceptualization of SES, the level of students' economic, social, and cultural possessions. This section focuses on socio-economic differences between Kazakhstani students that affect their reading literacy. A discussion on individual level SES differences is followed by a discussion on the school-level SES composition.

5.3.1 Student-Level SES Inequalities

Student and school SES were consistently found to be significant predictors of students' academic success across many countries (OECD, 2019c). Consistent with the previous studies, for Kazakhstani 15-year-olds, the multilevel analyses show that higher SES students achieve higher PISA scores compared to their lower SES peers. The effect of SES on reading literacy was found to be just over 5 points in PISA. This finding shows the presence of educational inequalities between Kazakhstani students based on their socio-economic standing in the society. It is important because one of the roles of the education system is to mitigate these SES differences and maintain social mobility of the population.

There are a number of plausible reasons that may explain the existence of achievement gaps in reading literacy based on students' SES found among Kazakhstani students. First, analyses show that reading literacy is particularly sensitive to students' SES, especially compared to math and science achievements. Further multilevel analysis of math and science literacies of Kazakhstani students showed a 3 and 2 points statistically significant positive effect, respectively. Reading requires physical educational resources not only at schools but also at home, which inevitably leads to family SES becoming a significant factor in developing students' reading motivation compared to STEM subjects. According to Belibas (2016), among the lowest SES students in Turkey, the unavailability of home educational

resources and ICT was the biggest challenge that increased the achievement gap. A similar situation exists in Kazakhstan, where in rural areas not only families but also schools lack educational resources and ICT. Second, as Neumann's (2016) and Hemmereichs, Agirdag, and Kavadias's (2017) studies suggests, achievement gaps in reading begin in early childhood based on parental involvement differences of different SES families. Lower SES families tend to have less involvement in children's reading activities at home, which appears to have a great impact on the later reading achievements of these students.

5.3.2 School-Level SES Inequalities

In addition to individual SES differences that account for reading literacy gaps among Kazakhstani students, the school mean SES was found to be a significant predictor of students' reading literacy success. The multilevel analysis findings show a 46-point advantage of students attending higher mean SES schools in PISA 2018 results. This finding once again refers to the presence of educational inequalities ignited by the prevalence of a certain group of students with similar SES in a school. It is important because not only does students' individual SES impact their educational achievements but also schools they attend play a major role in predicting their academic success, which contradicts the idea of fairness and inclusion in the education system. Interestingly, the same is true in the case of the math and science achievements of Kazakhstani students in PISA 2018, where multilevel analysis revealed a 48 and 40-point advantage of attending a higher mean SES school, respectively.

Although higher mean SES schools are usually associated with private and public schools, the findings suggest that school ownership is not a statistically significant predictor for the reading literacy performance of Kazakhstani students. On the other hand, math and science analyses show the statistical significance of the advantage influenced by the school ownership; in both cases, private school students outperforming public school students by 20 and 13-point difference. However, this is likely to be due to the school mean SES predictor

already accounting for the school ownership differences. Thus, it is to some extent safe to assume that school ownership seems to play a role in predicting students' PISA scores.

Three plausible reasons may explain the advantage of attending a higher mean SES schools in Kazakhstani context. First, higher mean SES schools are more likely to be located in resource-rich urban areas. Second, this gap may also be associated with an abundance of school resources in higher mean SES schools compared to lower mean SES schools. Third, these may be explained by the advantage of attending higher SES schools by disadvantaged students, which has a positive effect on their academic achievements, as suggested in studies of UK and US disadvantaged students in elite schools by Foliano et al. (2019) and Shi (2020), respectively.

Overall, it appears that disadvantaged students face a double risk: 1) initial disadvantages due to their family background; and 2) a higher probability of attending lower SES schools, where they might be exposed to a school population that has a negative peer effect. Thus, it is important to promote social inclusion in the education system.

5.4 School Characteristics and Educational Inequalities

While students' gender and SES, as well as school composition based on gender and SES prevalence within schools are important factors to consider in equity and excellence discourse, other school characteristics might play an even more significant role in risking the fairness and inclusion of the education system. This section focuses on the discussion of these school characteristics that include school type, school ownership, school resources in terms of average class size, infrastructure, and staff shortage.

The multilevel analysis revealed that school type is a significant predictor of students' reading literacy performance with by far the greatest effect on student performance equivalent to 76 points in PISA scores. This is the difference between students attending intellectual schools and the students attending non-intellectual schools, which is comparable to more than

three years of schooling difference. This finding refers to the substantial gap created by school types in Kazakhstan, which is important in consideration of equity within the system. On the one hand, intellectual schools are established to cater the needs of gifted students by competitive admission process and a more intensive curriculum. Thus, their advantage over the non-intellectual schools is to be expected. On the other hand, the limited number of these schools presents a problem, whereas other potentially gifted students who attend non-intellectual schools are left to study the standard curriculum and their needs are not met. As Yakavets (2014) explains, the problem lies in the identification-led approach to identifying and schooling gifted students in Kazakhstan. This Soviet style of education system where students are classified as either gifted or non-gifted at some point in school and taught accordingly in separate schools poses a threat to the equitability of the system. Hence, Kazakhstan started reforming the education system towards inclusion in classrooms using the differentiation approach to also meet the needs of gifted students within non-intellectual schools. However, the reform took its full scale in 2018 with all mainstream schools adopting intellectual schools' curriculum, and changes are to be expected only in the long-term.

Another finding to note is that reading literacy was found to be less sensitive to the school type predictor compared to math and science literacies among Kazakhstani schools. Intellectual schools exhibited 84 and 88-point advantage in PISA scores for math and science literacies, respectively. This finding seems to be an expected outcome of the specialization of intellectual schools in Kazakhstan in STEM subjects. Nevertheless, we see that these top-performing students admitted to STEM focused schools are also high performers in reading literacy. These findings are in line with the conceptualization of giftedness by McInerney (2013) who states that gifted students have above-average capabilities in various aspects and have an intrinsic motivation to learn. Thus, gifted students excelling in all PISA subject areas could be a topic of further research into giftedness among Kazakhstani students.

Interestingly, while school type was found to be an important predictor of students' academic success in this study, all other predictors describing school characteristics were found to be not statistically significant in predicting Kazakhstani students' reading literacy performance. Namely, the small or large classroom sizes, private or public schools, sufficient or poor school infrastructure, and high or low teaching staff shortage were found to have no significant effect on the reading literacy of students. These findings refer to school resources having no substantial effect on student excellence, which is important because lack of school resources is often associated with lack of equity and poor student performance. Hence, the findings of the multilevel analysis for Kazakhstani students contradict many studies that found school resources to be significant predictors of students' educational outcomes. There are two possible reasons behind these findings. First, other predictors could have accounted for these differences such as school mean SES and school types variables. In this case, higher mean SES schools and intellectual schools are more likely to be the ones with small class sizes, excellent school infrastructure, low teacher shortage, as well as high SES composition of schools similar to private schools. Second, these findings may also be explained by family SES being a more important predictor than school resources. As suggested by Hofflinger and Von Hippel (2020), parents becoming more educated and earning higher incomes had a greater impact on Chilean students' academic progress than an increase in school's educational resources.

5.5 Language Inequalities

In the previous section, the variables defining school characteristics were discussed in detail. This section continues the discussion of school level predictors effect on Kazakhstani students' reading literacy. Specific to the Kazakhstani context, the language of the assessment has been analyzed and the findings suggest that schools' language composition plays a major role in predicting students' reading literacy scores in PISA 2018.

The bilingual education system in Kazakhstan resulted in significant differences in students' reading abilities between Kazakh and Russian as a medium of instruction in schools. The gap accounted for nearly 76 points higher PISA results for students who passed the test in the Russian language. This finding is important as it implies the possible existence of educational inequalities associated with the language of study at schools. The advantage of Russian-speaking schools is also demonstrated by the language of the assessment being a significant predictor of math and science literacy performance of Kazakhstani students. Compared to 76 points effect on reading literacy, the Russian language of the assessment had 38 points and 66 points positive effect on math and science literacies, respectively. Math performance is less sensitive to the language of the assessment while reading literacy is the most sensitive to the language variable. Nevertheless, all three analyses show that schools with predominantly Russian speakers considerably outperform schools with predominantly Kazakh speakers. At first glance, the effect of language on the reading performance of students is to be expected, but the same effect being present for numeric subjects such as math and science is surprising. These findings indicate that the reason behind these significant differences caused by the language of the assessment could be more systematic in nature.

The most important reason that may explain this gap may be associated with the location of Russian and Kazakh as a medium of instruction in schools. Specific to the Kazakhstani context, schools that teach in the Russian language are mainly located in urban areas that boast higher mean SES and better school resources than rural areas. Thus, the achievement gap between the two language speakers is expected, nevertheless it is important for the government to improve rural schools to provide equal opportunities for learning.

5.6 Conclusion

The purpose of this chapter was to discuss the main findings of the multilevel analysis based on PISA data for Kazakhstani 15-year-old students. As seen above, the chapter began

by discussing the effect of gender and student SES on the reading literacy of students. The study findings were consistent with previous studies that found that girls outperform boys in reading literacy and higher SES students tend to perform better than their lower SES peers. Next, school characteristics represented by the school type, school ownership, school resources as the average class size, school infrastructure, and teacher shortage were discussed in relation to the students reading literacy performance in PISA. Contrary to many studies, school ownership and school resources were not significant predictors of Kazakhstani students' reading literacy. These findings could be attributable to the major educational differences created by the school types, which had one of the greatest effects on students' performance. Unique to Kazakhstani context, the existence of intellectual schools that benefit from competitive admission, better financing and governing, and abundant school resources contributed to the major inequitable treatment issues in the education system. Finally, the school language effect on students reading was discussed and possible reasons were offered to understand the meaning of the multilevel analyses' findings. As an officially bilingual country, Kazakhstani education system is unique in providing the assessment in two languages. The substantial advantage of Russian-speaking students in reading literacy highlights the issue of the language becoming one of the characteristics of disadvantaged students in Kazakhstan. The findings discussed in this chapter inform the further discussion of implications and future research in the next chapter.

Chapter 6. Conclusion

6.1 Introduction

This chapter summarizes and presents the final conclusions of this thesis. The purpose of this quantitative study was to evaluate the effects of student and school level factors on Kazakhstani students' reading literacy performance in PISA 2018. The study was guided by the main research question of how schools in Kazakhstan vary in their students' mean reading literacy scores in PISA 2018. The sub-questions included the research questions of 1) how students and schools' composition (i.e. gender, SES, language) affect students' reading literacy; 2) how school type and school ownership affected students' reading literacy; and 3) how school resources (i.e., average class size, school infrastructure, and availability of teachers) influence students' reading literacy. Answers to these research questions are discussed in detail in the first section of the chapter. In the next section, the implications for theory and practice of the findings of this thesis are addressed. The third section focuses on the limitations of this study and future research recommendations. The final section concludes the chapter and reiterates the most relevant findings and implications to close this thesis.

6.2 Revisiting Research Questions

This study aimed at measuring the effects of different factors on students' reading literacy in Kazakhstan. The findings from this study allow to respond to the research questions that guided this study as follows.

RQ 1: How do students and schools' composition (i.e. gender, SES, language) affect students' reading literacy in Kazakhstan?

The findings of this study suggest that students' gender and SES have significant effects on Kazakhstani students' reading literacy performance in PISA. In addition to that, it was found that schools' mean gender, SES, and language also exhibited a great effect on the students' reading literacy. A conclusion to be drawn from these findings is that Kazakhstani

students are not immune to gender and SES differences in terms of reading literacy and social inequalities inherently exist in the education system.

RQ 2: How do school type and school ownership affect the students' reading literacy?

A major finding of the study was that schools significantly differ in providing quality education to students in Kazakhstan. The results indicated that school type had one of the biggest impacts on predicting students' not only students' reading literacy, but also their math and science literacies in PISA 2018. However, school ownership differences showed no substantial effect on academic achievements. A conclusion emerging from these findings is that schools in Kazakhstan are not equitable and considerably contribute to the existing inequalities in the education system. A related conclusion is that school segregation matters more for schools that have competitive admission process rather than private or public schools. Thus, studying in private schools does not yield the same benefits in terms of academic achievements in reading literacy as selective schools based on academic merits.

RQ 3: How do school resources (i.e., average class size, school infrastructure, and availability of teachers) influence students' reading literacy?

Findings suggest that none of these qualities of school resources have significant effect on Kazakhstani students' reading literacy performance. A conclusion to be drawn from these findings is that access to well-resourced schools is not as important as other more significant factors that affect students' educational outcomes. In this regard, it can also be concluded that school resources do not seem to be contributing to existing educational inequalities, and students from well-resourced and poorly resourced schools perform at similar levels.

Main research question: What is the relationship between equity and excellence in Kazakhstani education system with regards to students' reading literacy?

After responding to the sub-questions of this research, the main research question can now be addressed. The main finding of this study was that schools in Kazakhstan vary in their

mean reading literacy scores mainly based on their type (intellectual or non-intellectual) and do not vary significantly based on school ownership and school resources. In addition to that, it was found that students' gender, SES, and language of study substantially contribute to these school mean reading literacy differences. A conclusion emerging from these findings is that the Kazakhstani education system seems to fail in mitigating social inequalities and contribute to increasing the achievement gap by creating a selective education system, where female, high-achieving, higher SES, and Russian-speaking students are favored compared to their peers.

6.3 Implications for Theory and Practice

This study offers interesting insights for the understanding of social inequalities in education in a post-Soviet Central Asian country. The findings from the present study strongly support the theory-based predictions suggested by western studies on equity and excellence, which argued that family background is a more important predictor of students' academic success than school resources. However, this study expanded the theory on the factors affecting students in a post-Soviet country by showing that unique to these states, the factors of school type and the language of study have an even greater effect on students' reading literacy. The existence of selective schools for gifted students and the dominance of the Russian language in schools seem to increase the achievement gap between students in Kazakhstan, thus deepening the social inequalities among them.

In practice, the findings of the study are difficult to be addressed as demonstrated by the aftermath of the Coleman Report. According to Hanushek (2016), half a century since the publication of the Report, the issue of family background and school SES composition predicting student achievement have not found a solution. However, in the case of Kazakhstan, the effect of school type and the language of study are far more important in predicting student achievement, each showing a three year of schooling gap in reading literacy

of 15-year-olds. Hence, a policy aimed at mitigating these differences would contribute towards significantly narrowing the gap not only in reading literacy, but also in math and science. Practicing differentiation in non-intellectual schools to meet the needs of diverse student population in these schools and increasing the quality of Kazakh as a medium of instruction schools, especially in rural areas, are the possible ways of improving student excellence. Both solutions require teacher professionalism in differentiated classrooms and providing quality teaching in remote areas as well. As Taole (2018) indicates, inclusive pedagogies and diverse instructional strategies are the priorities in improving student learning, especially in rural multigraded schools. Therefore, Kazakhstani policymakers should invest in quality pre-service and in-service teacher training to progress in tackling both equity and excellence issues in the education system.

6.4 Limitations and Future Research Recommendations

The results of the study should be interpreted in light of several limitations. First, the methodology of the study did not allow the use of schools which had less than 20 students participating in PISA 2018. This exclusion could potentially have impacted the findings since small size or ungraded schools usually have a very small number of 15-year-olds who have participated in the assessment. In other words, a specific disadvantaged group of students and schools could have been excluded from this study and affected the main findings by overestimating the regression coefficients and statistical significance of some factors. Second, although multilevel analysis requires school level predictors to vary not only between schools but also within schools, the language of assessment variable could be under/overestimated. While it was relatively easy to exclude single-gender schools due to their lower number, the exclusion of monolingual schools could have undermined the statistical rigor of the study due to their large number. Therefore, monolingual schools were included in the study and could have impacted the analysis and results of the study. Third, it is also necessary to acknowledge

that despite data being drawn from an established organization and rigorous assessment, the questionnaires on school and student characteristics were completed by students and school administration, which could potentially result in some inconsistent and inaccurate data. For instance, there were cases when similar NIS schools reported to be private and public at the same time, which is impossible since all NIS schools are structured and organized in the same way. Hence, it is important to be careful in interpreting the findings of this study and PISA results in general.

In particular, it is important to recognize PISA limitations in its reliability and validity. Schleicher (2017) points out there are concerns in the relevance of the curriculum for designing items in PISA and also cultural differences that should not be overlooked in cross-national assessments. Hopfenbeck (2016) identifies more specific issues related to PISA, namely 1) different language demands due to translations of items; 2) the need for secondary analysis of PISA data due to errors in sampling or interpreting the findings; 3) the level of reliability of self-report by students and school administration in PISA questionnaires; and 4) the significant differences in digital and print reading.

Despite these limitations, this study uses a quantitative approach using multilevel regression analyses that offer a novel approach and meaningful contribution to the study of equity and excellence in reading literacy based on PISA data. Thus, the findings of this study might be useful for policymakers who needs a study of measurable effects on students' academic achievements performed on a large dataset from a credible source.

Based on the findings of the study and the limitations indicated above, some recommendations for future research can be outlined. For example, future studies could replicate the findings of this study by including the sample of schools that have less than 20 students participating in PISA and perform additional statistical analysis of the significance of the findings. Also, other studies should consider exploring the equity and excellence in

Kazakhstan by examining similar variables but using only one language of study and determine the significance of the student and school-level factors separately for Kazakh as a medium of instruction schools and Russian as a medium of instruction schools. Additionally, further research could investigate the location of schools. This was not possible in this study due to unavailability of data from OECD on urban and rural location of schools in Kazakhstan, and the study of regional differences of schools required more advanced statistical techniques that was beyond the scope of the current study.

6.5 Conclusion

The present study highlights the importance of promoting equity in the education system to ensure a high-quality education system that results in student excellence. Overall, the study suggests that student gender, SES, language of study, and the school type were significant predictors of Kazakhstani students' reading literacy performance in PISA 2018. Furthermore, the findings suggest that school ownership and school resources are not statistically significant predictors of students' reading literacy. These findings are to some extent consistent with previous studies that found that girls, higher SES students, and school types are important factors for predicting reading achievements. In addition, this study found that language of study is a significant factor for predicting students' reading literacy in the plurilingual society. Contrary to previous studies, the current study also established that the ownership of the school (private vs public), and the availability of school resources are not significant predictors of students' reading literacy. The implications of these findings for theory of studying equity and excellence in a post-Soviet country included the effects of school type and language of study on top of family background and school SES predictors. Implications for practice are the importance of cultivating teacher professionalism for differentiated classrooms in non-intellectual schools and quality teaching in Kazakh as a medium of instruction schools, especially in rural areas of the country.

References

- Abad, D., & Cabrera-Alvarez, P. (2020). Bilingual autonomous communities, identities and educational achievement according to PISA 2015. *Revista de Educacion*, 2020(387), 157-180. DOI: 10.4438/1988-592X-RE-2020-387-443
- Agasisti, T., & Longobardi, S. (2014). Inequality in education: Can Italian disadvantaged students close the gap? *Journal of Behavioral and Experimental Economics*, 52, 8-20. DOI: 10.1016/j.socec.2014.05.002
- Alegre, M.A., & Ferrer, G. (2010). School regimes and education equity: Some insights based on PISA-2006. *British Educational Research Journal*, 36(3), 433-461. DOI: 10.1080/01411920902989193
- Areenpattamannil, S., & Kaur, B. (2013). Factors predicting science achievement of immigrant and non-immigrant students: A multilevel analysis. *International Journal of Science and Mathematics Education*, 11(5), 1183-1207. DOI: 10.1007/s10763-012-9369-5
- Becker, M., & McElvany, N. (2018). The interplay of gender and social background: A longitudinal study of interaction effects in reading attitudes and behaviour. *British Journal of Educational Psychology*, 88(4), 529-549. DOI: 10.1111/bjep.12199
- Belibas, M.S. (2016). Who are the most disadvantaged? Factors associated with the achievement of students with low socio-economic backgrounds. *Educational Sciences: Theory and Practice*, 16(2), 691-710. DOI: 10.12738/estp.2016.2.0257
- Bodovski, K., Jeon, H., & Byun, S. (2017). Cultural capital and academic achievement in post-socialist Eastern Europe. *British Journal of Sociology of Education*, 38(6), 887-907. DOI: 10.1080/01425692.2016.1202746
- Chapman, S. (2020). The significance of context: Autonomy and curriculum reform in rural schools. *Curriculum Journal*, 31(2), 1-13. DOI: 10.1002/curj.27

- Chen, Q. (2016). A multilevel analysis of Singaporean students' mathematics performance in PISA 2012. In L.M. Thien, N.A. Razak, J.P. Keeves, & I.G.N. Darmawan (Eds.), *What can PISA 2012 data tell us?* (pp.17-33). Rotterdam: Sense Publishers. DOI: 10.1007/978-94-6300-468-8_2
- Chiu, M. (2015). Family inequality, school inequalities, and mathematics achievement in 65 countries: Microeconomic mechanisms of rent seeking and diminishing marginal returns. *Teachers College Record*, 117(1). Retrieved from https://www.researchgate.net/publication/281740451_Family_inequality_school_inequalities_and_mathematics_achievement_in_65_countries_Microeconomic_mechanisms_of_rent_seeking_and_diminishing_marginal_returns
- Cobb-Clark, D., & Moschion, J. (2017). Gender gaps in early educational achievement. *Journal of Population Economics*, 30(4), 1093-1134. DOI: 10.1007/s00148-017-0638-z
- Creswell, J. W. (2014). *Educational research: Planning, conducting and evaluating quantitative and qualitative research (4th ed.)*. Harlow, Essex: Pearson.
- De Jager, T. (2017). Perspectives of teachers on differentiated teaching in multi-cultural South African secondary schools. *Studies in Educational Evaluation*, 53, 115-121. DOI: 10.1016/j.stueduc.2016.08.004
- De Naegel, J., Valcke, M., De Meyer, I., Warlop, N., Van Braak, J., & Van Keer, H. (2014). The role of teacher behaviour in adolescents' intrinsic reading motivation. *Reading and Writing*, 27(9), 1547-1565. DOI: 10.1007/s11145-014-9506-3
- Dickinson, E.E. (2016). Coleman Report set the standard for the study of public education. *Johns Hopkins Magazine*. Retrieved from <https://hub.jhu.edu/magazine/2016/winter/coleman-report-public-education/>

- Dolean, D., Melby-Lervag, M., Tincas, I., Damsa, C., & Lervag, A. (2019). Achievement gap: Socioeconomic status affects reading development beyond language and cognition in children facing poverty. *Learning and Instruction, 63*(2019), 1-10. DOI: 10.1016/j.learninstruc.2019.101218
- Elliott, L., & Bachman, H. (2018). Parents' educational beliefs and children's early academics: Examining the role of SES. *Children and Youth Services Review, 91*, 11-21. DOI: 10.1016/j.chilyouth.2018.05.022
- Evans, S. (2020). When schools choose: Evaluation and educational inequality. *Sociological Quarterly, 1*-21. DOI: 10.1080/00380253.2019.1711248
- Foliano, F., Green, F., & Sartarelli, M. (2019). Away from home, better at school. The case of a British boarding school. *Economics of Education Review, 73*. DOI: 10.1016/j.econedurev.2019.101911
- Fredriksson, U., Holzer, T., McCluskey-Cavin, H., & Taube, K. (2009). Strengths and weaknesses in the Swedish and Swiss education systems: A comparative analysis based on PISA data. *European Educational Research Journal, 8*(1), 54-68. DOI: 10.2304/eej.2009.8.1.54
- Gilleece, L., Cosgrove, J., & Sofroniou, N. (2010). Equity in mathematics and science outcomes: Characteristics associated with high and low achievement on PISA 2006 in Ireland. *International Journal of Science and Mathematics Education, 8*(3), 475-496. DOI: 10.1007/s10763-010-9199-2
- Guba, E.G., & Lincoln, Y.S. (1994). Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.
- Hanushek, E.A. (2016). What matters for student achievement: Updating Coleman on the influence of families and schools. *Education Next, 16*(2), 23-30. Retrieved from

<http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202016%20EdNext%2016%282%29.pdf>

Hemmerechts, K., Agirdag, O., & Kavadias, D. (2017). The relationship between parental literacy involvement, socio-economic status and reading literacy. *Educational Review*, 69(1), 85-101. DOI: 10.1080/00131911.2016.1164667

Hernández-Torrano, D., Tursunbayeva, X., & Almukhambetova, A. (2019). Teachers' conceptions about giftedness and gifted education: An international perspective. In M. A. Zanetti, G. Gualdi, M. Cascianelli (eds.), *Understanding Giftedness: A guide for parents and educators*. London: Routledge.

Hofflinger A., & Von Hippel, P. (2020). Does achievement rise fastest with school choice, school resources, or family resources? Chile from 2002 to 2013. *Sociology of Education*, 2020, 1-54. DOI: 10.1177/0038040719899358

Holt, L., & Bowlby, S. (2019). Gender, class, race, ethnicity and power in an elite girls' state school. *Geoforum*, 105, 168-178. DOI: 10.1016/j.geoforum.2019.05.020

Hopfenbeck, T.N. (2016). The power of PISA – limitations and possibilities for educational research. *Assessment in Education: Principles, Policy & Practice*, 23(4), 423-426. DOI: 10.1080/0969594X.2016.1247518

Hox, J.J., Moerbeek, M., & Schoot, R. (2018). *Multilevel analysis: Techniques and applications*. New York: Routledge.

Huang, J., Tang, Y., He, W., & Li, Q. (2019). Singapore's school excellence model and student learning: Evidence from PISA 2012 and TALIS 2013. *Asia Pacific Journal of Education*, 39(1), 96-112. DOI: 10.1080/02188791.2019.1575185

Information and Analytical Center. (2018). *Natsionalnyi Doklad o Sostoianii i Razvittii Sistemy Obrazovania Respubliki Kazakhstan (za gody Nezavisimosti Kazakhstana) [National Report on the State and Development of Education*

System of the Republic of Kazakhstan (During the Independence of Kazakhstan)].

Nur-Sultan: Information and Analytical Center. Retrieved from

http://iac.kz/sites/default/files/0_nacionalnyy_doklad_za_2018_god_final_s_oblozhkami_na_sayt_compressed_0.pdf

Jenkins, S.P., Micklewright, J., & Schnepf, S.V. (2008). Social segregation in secondary schools: How does England compare with other countries? *Oxford Review of Education*, 34(2008), 21-37. DOI: 10.1080/03054980701542039

Jerrim, J., & Sims, S. (2019). How do academically selective school systems affect pupils' social-emotional competencies? New evidence from the millennium cohort study. *American Educational Research Journal*, 56(5), 1769-1799. DOI: 10.3102/0002831219830965

Karakolidis, A., Pitsia, V., & Emvalotis, A. (2016). Examining students' achievement in mathematics: A multilevel analysis of the Programme for International Student Assessment (PISA) 2012 data for Greece. *International Journal for Educational Research*, 79, 106-115. DOI: 10.1016/j.ijer.2016.05.013

Kenayathulla, H., Ling, H., Razak, A., & Darusalam, G. (2019). School level resources and students' performance in Malaysian national type Chinese schools. *Malaysian Online Journal of Educational Management*, 7(1), 37-56. DOI: 10.22452/mojem/vol7no1.3

Maag Merki, K., Buehlmann, F., Kamm, C., Truniger, A., & Emmerich, M. (2020). Support of students in primary schools: A comparative case study in a selective education system. *Journal of Curriculum Studies*, 1-19. DOI: 10.1080/00220272.2020.1734663

- Mahuteau, S., & Mavromaras, K. (2014). An analysis of the impact of socio-economic disadvantage and school quality on the probability of school dropout. *Education Economics*, 22(4), 389-411. DOI: 10.1080/09645292.2014.918586
- McInerney, D.M. (2013). The gifted and talented and effective learning: A focus on the actiotope model of giftedness in the Asian context. In S.N. Phillipson, H. Stoeger, & A. Ziegler, *Exceptionality in East Asia: Explorations in the actiotope model of giftedness* (pp.255-268). London: Routledge.
- Mclaughlin, C. (2017). Researching practice as education and reform. In M.A. Peters, C. Bronwen, & I. Menter (Ed.), *A companion to research in teacher education* (pp.583-597). Singapore: Springer Singapore.
- Muelle, L. (2020). Socioeconomic and contextual factors associated with low academic performance of Peruvian students in PISA 2015. *Apuntes*, 47(86), 111-146. DOI: 10.21678/apuntes.86.943
- Ndethiu, S., Masingila, J., Miheso-O'Connor, M., Khatete, D., & Heath, K. (2017). Kenyan secondary teachers' and principals' perspectives and strategies on teaching and learning with large classes. *Africa Education Review*, 14(1), 58-86. DOI: 10.1080/18146627.2016.1224573
- Neumann, M. (2016). A socioeconomic comparison of emergent literacy and home literacy in Australian pre-schoolers. *European Early Childhood Educational Research Journal*, 24(4), 555-566. DOI: 10.1080/1350293X.2016.1189722
- Ning, B., Van Damme, J., Liu, H., Vanlaar, G., & Gielen, S. (2013). Students' individual perceptions of school climate predict reading achievement in Shanghai. *International Journal of School and Educational Psychology*, 1(3), 188-198. DOI: 10.1080/21683603.2013.821638

O'Connell, M. (2019). Is the impact of SES on educational performance overestimated?

Evidence from the PISA survey. *Intelligence*, 75(2019), 41-47. DOI:

10.1016/j.intell.2019.04.005

Oberti, M., & Savina, Y. (2019). Urban and school segregation in Paris: The complexity of

contextual effects on school achievement: The case of middle schools in Paris

metropolitan area. *Urban Studies*, 56(15), 3117-3142. DOI:

10.1177/0042098018811733

OECD. (2009). *PISA data analysis manual*. Paris: OECD Publishing.

OECD. (2010). *The high cost of low educational performance: The long-run economic*

impact of improving PISA outcomes. Paris: OECD Publishing.

OECD. (2014). *Equity, excellence and inclusiveness in education: Policy lessons from*

around the world. Paris: OECD Publishing.

OECD. (2015). *OECD Reviews of School Resources: Kazakhstan 2015*. Paris: OECD

Publishing.

OECD. (2019a). *PISA 2018 results (Volume I): What students know and can do*. Paris:

OECD Publishing.

OECD. (2019b). *PISA 2018 results (Volume II): Where all students can succeed*. Paris:

OECD Publishing.

OECD. (2019c). *PISA 2018 results: Combined executive summaries*. Paris: OECD

Publishing.

Oliden, P., & Lizaso, J. (2014). Impact of family language and testing language on reading

performance in bilingual educational context. *Psicothema*, 26(3), 328-335. DOI:

10.7334/psicothema2013.344

Ozdemir, C. (2016). Equity in the Turkish education system: A multilevel analysis of

social background influences on the mathematics performance of 15-year-old

students. *European Educational Research Journal*, 15(2), 193-217. DOI:
10.1177/1474904115627159

PISA 2018 Database. (n.d.). Retrieved December 15, 2019, from

<https://www.oecd.org/pisa/data/2018database/>

Pitsia, V., Biggart, A., & Karakolidis, A. (2017). The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement: A multilevel analysis of the Programme for International Student Assessment data. *Learning and Individual Differences*, 55, 163-173. DOI: 10.1016/j.lindif.2017.03.014

Qin, L., & Bowen, D. (2019). The distributions of teacher qualifications: A cross-national study. *International Journal of Educational Development*, 70. DOI:
10.1016/j.ijedudev.2019.102084

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and data analysis methods (2nd ed.)*. Thousand Oaks: SAGE Publications.

Rogiers, A., Van Keer, H., & Merchie, E. (2020). The profile of the skilled reader: An investigation into the role of reading enjoyment and student characteristics. *International Journal of Educational Research*, 99, 1-14. DOI:
10.1016/j.ijer.2019.101512

Sakellariou, C. (2017). Private or public-school advantage? Evidence from 40 countries using PISA 2012 Mathematics. *Applied Economics*, 49(29), 1-18. DOI:
10.1080/00036846.2016.1248361

Scherbaum, C.A., & Ferrerter, J.M. (2009). Estimating statistical power and required sample size for organizational research using multilevel modelling. *Organizational Research Methods*, 12(2), 347-367. DOI:
10.1177/1094428107308906

- Schleicher, A. (2017). Seeing education through the prism of PISA. *European Journal of Education, 2017*(52), 124-130. DOI: 10.1111/ejed.12209
- Schmidt, W.H., & Burroughs, N.A. (2016). The trade-off between excellence and equality: What international assessments tell us. *Georgetown Journal of International Affairs, 17*(1), 103-109. DOI: 10.1353/gia.2016.0011
- Shera, P. (2014). School effects, gender and socioeconomic differences in reading performance: A multilevel analysis. *International Education Studies, 7*(11), 28-39. DOI: 10.5539/ies.v7n11p28
- Shi, Y. (2020). Who benefits from selective education? Evidence from elite boarding school admissions. *Economics of Education Review, 74*(2020), 1-26. DOI: 10.1016/j.econedurev.2019.07.001
- Sierra, J. (2008). Assessment of bilingual education in the Basque Country. *Language, Culture and Curriculum, 21*(1), 39-47. DOI: 10.2167/lcc341.0
- Skutnabb-Kangas, T. (1981). *Bilingualism or not: The education of minorities (3rd ed.)*. Avon, England: Multilingual Matters Ltd.
- State Programme of Education and Science Development 2016-2019. (2016). Decree of the President of Kazakhstan No.205. Nur-Sultan: “Akorda” Presidential Residence. Retrieved from <https://bit.ly/2BkfDOG>
- Tan, C., & Hairon, S. (2016). Education reform in China: Toward classroom communities. *Action in Teacher Education, 38*(4), 315-326. DOI: 10.1080/01626620.2016.1226205
- Taole, M. (2018). Diversity and inclusion in rural South African multigrade classrooms. *International Journal of Inclusive Education, 2018*, 1-17. DOI: 10.1080/13603116.2018.1520310

- Thien, L. (2016). Malaysian students' performance in Mathematics literacy in PISA from gender and socioeconomic status perspectives. *Asia-Pacific Education Researcher*, 25(4), 657-666. DOI: 10.1007/s40299-016-0295-0
- Topcu, M., Arikan, S., & Erbilgin, E. (2015). Turkish students' science performance and related factors in PISA 2006 and 2009. *Australian Educational Researcher*, 42(1), 117-132. DOI: 10.1007/s13384-014-0157-9
- Trinidad, J. (2020). Material resources, school climate, and achievement variations in the Philippines: Insights from PISA 2018. *International Journal of Educational Development*, 75(2020),1-37, DOI: 10.1016/j.ijedudev.2020.102174
- Tsai, S.L., Smith, M.L., & Hauser, R.M. (2017). Families, schools, and student achievement inequality: A multilevel MIMIC model approach. *Sociology of Education*, 90(1), 64-88. DOI: 10.1177/0038040716683779
- 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. DOI: 10.1080/13803610600587016
- Yakavets, N. (2014). Reforming society through education for gifted children: The case of Kazakhstan. *Research Papers in Education*, 29(5), 513-533. DOI: 10.1080/02671522.2013.825311

Appendices

Appendix A

Multilevel Models with Main Effects for Math Literacy

	Model 0	Model 1	Model 2	Model 3
	Intercept-only	Within-school effects model	Between-school effects model	Full model
Intercept	427.7 (2.0)	421.7 (3.0)	472.3 (9.4)	460.8 (20.9)
<i>Student-level</i>				
Gender (Male)		5.7 (1.1)**	5.9 (1.1)**	5.9 (1.1)**
ESCS		4.0 (0.8)**	2.9 (0.8)**	2.9 (0.8)**
<i>School-level</i>				
Clsize			-0.2 (0.1)	0.1 (0.1)
Staffshort			0.4 (2.1)	-0.4 (1.9)
Edushort			-1.8 (2.0)	-1.8 (1.9)
Sch_int (Intellectual)			79.6 (9.7)**	84.4 (9.0)**
Sch_own (Public)			-25.4 (6.8)**	-20.0 (6.3)**
ESCS.sch			54.2 (5.0)**	48.2 (4.9)**
Language.sch				38.2 (4.7)**
Gender.sch				-14.5 (12.8)
Units (school/student)	432/17820	432/17820	432/17820	432/17820
-2*log-likelihood	-102269.1	-102143.7	-102095.1	-102090.8

Note. * $p < .05$; ** $p < .01$; -2*log-likelihood indicates a relative fit of the model to the data.

Appendix B

Multilevel Models with Main Effects for Science Literacy

	Model 0	Model 1	Model 2	Model 3
	Intercept-only	Within-school effects model	Between-school effects model	Full model
Intercept	400.6 (1.9)	410.8 (2.8)	467.5 (8.8)	440.7 (15.5)
<i>Student-level</i>				
Gender (Male)		-4.4 (1.0)**	-4.3 (1.0)**	-4.2 (1.0)**
ESCS		3.3 (0.6)**	2.4 (0.6)**	2.4 (0.6)**
<i>School-level</i>				
Clsize			-0.6 (0.1)**	-0.03 (0.1)
Staffshort			-0.5 (2.0)	-1.8 (1.4)
Edushort			2.0 (1.9)	2.0 (1.4)
Sch_int (Intellectual)			79.9 (9.2)**	88.3 (6.6)**
Sch_own (Public)			-22.0 (6.4)**	-12.6 (4.7)**
ESCS.sch			50.9 (4.7)**	40.9 (3.6)**
Language.sch				66.5 (3.5)**
Gender.sch				-20.7 (9.5)*
Units (school/student)	432/17820	432/17820	432/17820	432/17820
-2*log-likelihood	-99484.2	-99451.3	-99297.8	-99293.1

Note. * $p < .05$; ** $p < .01$; -2*log-likelihood indicates a relative fit of the model to the data.