The Influence of School Climate on Kazakhstani Students' Academic Achievements in the Programme for International Student Assessment (PISA)

Akzhol Akhmetov

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

May 7, 2021

Word Count: 17,329

AUTHOR AGREEMENT

Declaration

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: Akzhol Akhmetov

Date: 07.05.2021

Ethical Approval

NAZARBAYEV UNIVERSITY
Graduate School of Education

53 Kabanbay Batyr Ave.
010000 Astana,
Republic of Kazakhstan
October 2020

Dear Akzhol Akhmetov,

This letter now confirms that your research project entitled: The Influence of School Climate on Kazakhstani Students' Academic Achievements in the Programme for International Student Assessment (PISA) has been approved by the Graduate School of Education Ethics Committee of Nazarbayev University.

You may proceed with contacting your preferred research site and commencing your participant recruitment strategy.

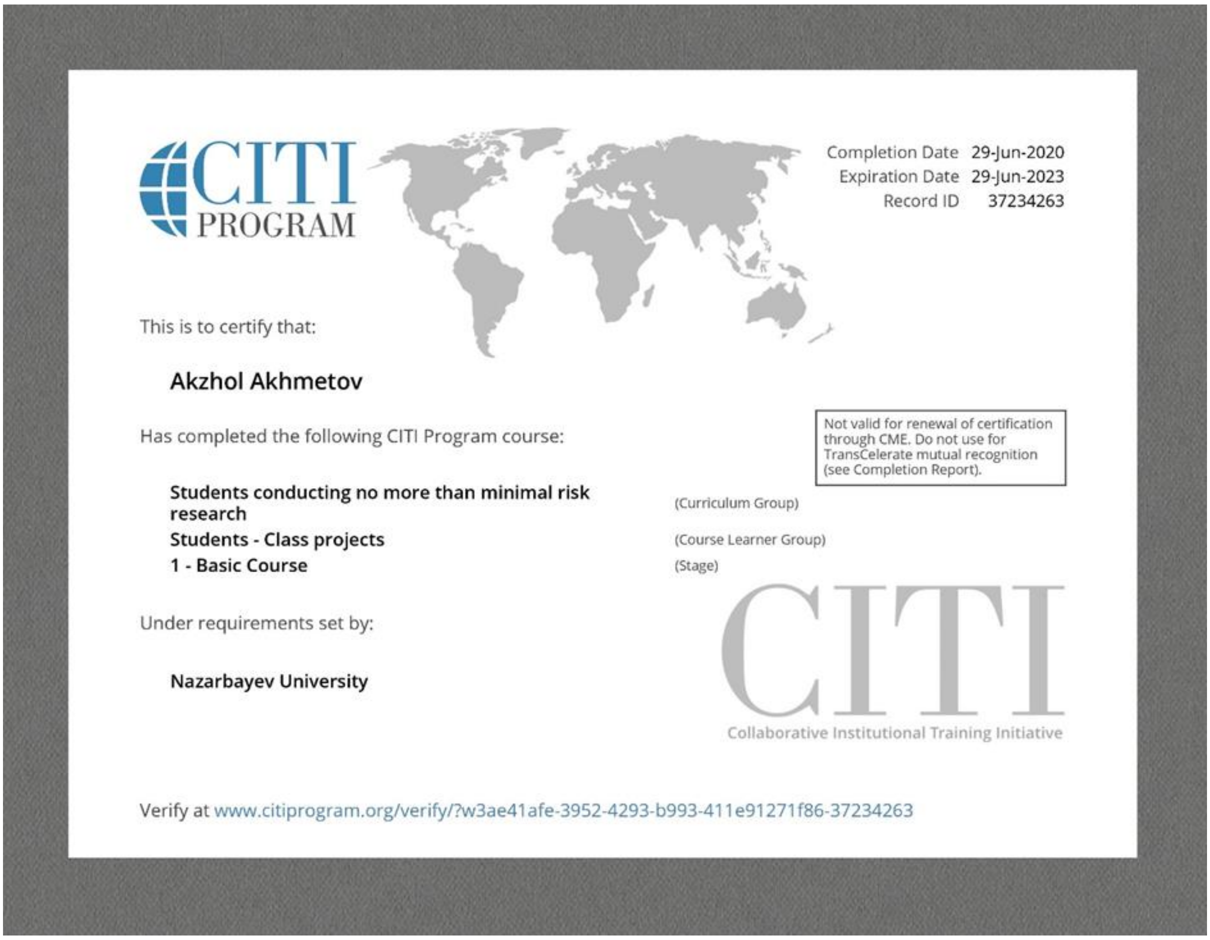
Yours sincerely,

Assistant Professor Matthew Courtney, PhD

On behalf of Zumrad Kataeva
Chair of the GSE Ethics Committee
Assistant Professor
Graduate School of Education
Nazarbayev University

Block C3, Room 5006
Office: +7 (7172) 70 9371
Mobile: +7 777 1929961
email: zumrad.kataeva@nu.edu.kz

CITI training certificate



CITI PROGRAM

Completion Date 29-Jun-2020
Expiration Date 29-Jun-2023
Record ID 37234263

This is to certify that:

Akzhol Akhmetov

Has completed the following CITI Program course:

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

Students conducting no more than minimal risk research (Curriculum Group)
Students - Class projects (Course Learner Group)
1 - Basic Course (Stage)

Under requirements set by:

Nazarbayev University

CITI
Collaborative Institutional Training Initiative

Verify at www.citiprogram.org/verify/?w3ae41afe-3952-4293-b993-411e91271f86-37234263

Acknowledgment

I would like to thank all NUGSE professors and staff for their assistance and shared experience. My sincere gratitude to my supervisor Dr. Matthew Courtney, for guiding and supporting us throughout the whole thesis journey. Many thanks for his complete dedication!

The Influence of School Climate on Kazakhstani Students' Academic Achievements in the Programme for International Student Assessment (PISA)

Abstract

The importance of school climate for students' academic achievements has been supported by many studies. However, there has been done little to no research in this area in Kazakhstan. Therefore, it was crucial to study this topic. The purpose of this quantitative study is to identify the influence of school climate-related factors on Kazakhstani students' academic outcomes in PISA Math, Reading, and Science literacy. In the study, there was no sufficient between-school variation in school climate. Accordingly, the appropriate method defaulted to a single-level form of multivariate analysis. Ten school climate factors including, an Anti-bullying attitude, the Experience of bullying, the Disciplinary climate, Teacher enthusiasm, Support, and Behavior, Student cooperation and competition, a Lack of sense of belonging to school, and Parental involvement in school activities were measured. The findings of this study both supported and disproved previous studies. For example, only two out of ten school climate factors were the main predictors of student achievement in all three subjects of interest. The research has an Anti-bullying attitude as the main predictor of student performance. The second most important factor for student performance in all disciplines was the Experience of bullying. However, most of the factors almost did not affect PISA student performance despite being statistically significant. Out of all categories, three factors, Teacher behavior, Student cooperation, and Teacher support were found insignificant at least to one of the outcomes of interest. However, Teacher behavior was the least important factor being insignificant for all three subjects. The findings of the research might be useful for different stakeholders, including principals, policymakers, and teachers looking to make changes to school policy and school climate with the primary intention of improving the academic outcomes of students.

Key words: school climate, PISA, academic achievements

Влияние школьного климата на академические достижения казахстанских студентов в Международной программе по оценке образовательных достижений учащихся (PISA)

Абстракт

Важность школьного климата для успеваемости учащихся подтверждается многими исследованиями. Однако в Казахстане практически не проводились исследования в этой области. Поэтому исследование данной темы было крайне важным. Целью этого количественного исследования является определение влияния факторов, связанных с климатом в школе, на успеваемость казахстанских учащихся по математике, чтению и естественным наукам в рамках программы PISA. В ходе исследования не было достаточных различий в школьном климате между образовательными учреждениями. Соответственно, подходящим методом по умолчанию использовалась одно уровневая форма многомерного анализа. Были измерены десять факторов школьного климата, в том числе Отношение к запугиванию, Подвергание запугиванию, Дисциплинарный климат, Энтузиазм, Поддержка и Поведение учителя, Сотрудничество и Конкуренция учеников, Отсутствие чувства принадлежности к школе и Участие родителей в школьной деятельности. Результаты этого исследования подтвердили и опровергли предыдущие исследования. Например, только два из десяти факторов школьного климата были основными предикторами успеваемости учащихся по всем трем интересующим предметам. Основным влияющим показателем успеваемости учащихся в исследовании является Отношение к запугиванию. Вторым по важности фактором успеваемости учащихся по всем дисциплинам был Подвергание запугиванию. Однако большинство факторов почти не повлияли на успеваемость учащихся PISA, несмотря на их статистическую значимость. Из всех категорий три фактора: Поведение учителя, Сотрудничество учеников и Поддержка учителя были признаны несущественными по крайней мере для одного из представляющих интерес результатов. Тем не менее Поведение учителя было наименее важным фактором, так как не имел существенной значимости для всех трех предметов. Результаты исследования могут быть полезны для различных заинтересованных сторон, включая директоров, политиков и учителей, стремящихся внести изменения в школьную политику и школьный климат с основной целью улучшить академические результаты учащихся.

Ключевые слова: школьный климат, PISA, академические достижения.

Мектеп климатының қазақстандық оқушылардың Халықаралық оқушы қабілетін бағалау бағдарламасындағы (PISA) оқу жетістіктеріне әсері

Аңдатпа

Оқушылардың жетістіктері үшін мектеп климатының маңыздылығын көптеген зерттеулер дәлелдейді. Алайда, Қазақстанда бұл бағытта іс жүзінде ешқандай зерттеулер жүргізілген жоқ. Сондықтан бұл тақырыпты зерттеу өте маңызды болды. Бұл сандық зерттеудің мақсаты мектеп климатына байланысты факторлардың қазақстандық оқушылардың PISA бағдарламасының математика, оқу және жаратылыстану пәндеріне әсерін анықтау болып табылады. Зерттеу барысында білім беру мекемелері арасындағы мектеп климатында жеткілікті айырмашылықтар болған жоқ. Тиісінше, әдепкі сәйкес әдіс ретінде көп деңгейлі талдаудың бір деңгейлі формасы қолданылды. Мектептегі климаттың он факторы өлшенді, оның ішінде Жәбірленушіге деген көзқарас, Жәбірлену тәжірибесі, Тәртіптік климат, Мұғалімнің Қолдауы, Ынта-жігері, мен Мінез-құлқы, Оқушылардың Ынтымақтастығы мен Бәсекелестігі, Мектепке деген сезімнің болмауы және Ата-аналардың мектеп іс-шараларына қатысуы тексерілген болатын. Осы зерттеудің нәтижелері алдыңғы зерттеулерді бір жағынан растаса екінші жағынан жоққа шығарды. Мысалы, мектеп климатындағы он фактордың тек екеуі ғана қызығушылық тудыратын барлық үш пән бойынша оқушылардың жетістіктерінің негізгі болжаушылары болды. Зерттеу барысында оқушылардың жұмысына әсер ететін негізгі көрсеткіш Жәбірленушіге деген көзқарас болып шықты. Барлық пәндер бойынша студенттер жетістіктерінің екінші маңызды факторы Жәбірлену тәжірибесі болды. Алайда, факторлардың көпшілігі олардың статистикалық маңыздылығына қарамастан студенттерінің PISA көрсеткіштерәне айтарлықтай әсер етпеді. Барлық санаттардың ішінде үш фактор, яғни Мұғалімдердің мінез-құлқы, Оқушылардың ынтымақтастығы және Мұғалім қолдауы, қызығушылық нәтижелерінің ең болмағанда біреуі үшін маңызды емес деп саналды. Алайда Мұғалімнің мінез-құлқы ең маңызды емес фактор болды, өйткені үш пән үшін де маңызды емес болып шықты. Зерттеулердің нәтижелері студенттердің оқу үлгерімін жақсарту мақсатында мектеп саясатына және мектеп климатына өзгерістер енгізуге ұмтылатын директорлар, саясаткерлер мен мұғалімдерді қоса алғанда әр түрлі мүдделі тараптарға пайдалы болуы мүмкін.

Түйінді сөздер: мектеп климаты, PISA, оқу жетістіктері.

Table of contents

List of tables

List of figures

1. Introduction

1.1 Background Information

1.1.1 School Climate

School climate has always been an essential part of education as it sets momentum in the learning process. A favorable school climate may provide students with more chances to succeed academically. For example, the importance of school climate is highlighted by research that suggests that the success of students in school depends on school ambiance (Deal & Peterson, 1998; Shindler, Jones, Williams, Taylor, & Cardenas, 2016).

According to the OECD (2019a) school climate can be categorized as “safe or unsafe, cohesive or divisive, collaborative or competitive” (p. 37). Furthermore, the school climate is classified as generally positive or negative. Research has paid attention to the importance of more positive school climates for student learning. The OECD (2019a) reports that positive school climate initiates certain beneficial conditions for all stakeholders. For example, in a favorable school ambience students feel emotional and physical safety; teachers are more helpful, passionate, and compassionate; parents willingly take part in school activities; school society is created around friendly and harmonious relationships; and all the stakeholders look after the school surrounding and work collaboratively to develop school community (p.37).

The Kazakhstani government is paying close attention to improving schools by increasing teacher status and providing conditions for teachers and students (Ministry of Education and Science of the Republic of Kazakhstan [MoES], 2016). However, the role of school climate is neglected in Kazakhstan. To the best of this author’s knowledge, no empirical work on the topic has been undertaken.

Several factors can affect the school climate, and it is quite clear that principals are one of the main agents in establishing a positive school atmosphere, so it is crucial to understand their role, attitude, and the conception that they have of their school and its environment. A principal’s role, attitude, and conception of a school’s climate are essential for encouraging the teaching staff and students and for promoting school improvement. In majority of schools, the school principals play a pivotal role in organizational processes (Frost, Fimyar, Yakavets, & Bilyalov, 2014). According to Frost et al. (2014), school leaders in Kazakhstan are central figures for stimulating teachers and students by creating a pleasant atmosphere and assisting to initiate fairness at school. Deal and Peterson (1998) also highlight the importance of principals as leaders and the major factor in establishing a school climate and making positive transformations.

The extent to which between-school variation in school climate factors exists in Kazakhstan is yet to be explored. Moreover, in order to integrate and measure the effect of school-level variables in this study, there must be sufficient variation in school climate variables at the between-school level. If there is not sufficient variation in between-school effects for school climate, then the analysis of school-level variables will be redundant. Therefore, the extent to which between-school differences exist in school climate in Kazakhstan will be explored in the first part of this study.

While school climate is important, students’ educational achievement should also be taken into account as it is a fundamental part of the educational domain. The focus here should not solely be on the improvement of students' test scores but the development of their functional literacy, numeracy, and problem-solving abilities. To develop students' skills and abilities, the Kazakhstani government implemented several reforms such as creating a network of twenty Nazarbayev Intellectual Schools (NIS) and the introduction of updated curriculum (MES, 2016; Muratkyzy, 2020, p. 21). Accordingly, to meet the standards of innovation in the modern world, the government introduced an updated curriculum and reforms associated with the assessment system (MES, 2016). However, to the best of this author’s knowledge, no research has been undertaken on how a school’s climate might contribute to improved student academic outcomes in this updated curricula and pedagogical context. Therefore, this research aims to analyze the role of school climate on the three main subject domains in PISA, namely Reading, Science, and Mathematics literacy. A description of these three outcome variables is now provided.

1.1.2 PISA Reading, Science, and Mathematics Literacy

As for reading literacy, according to PISA it can be described in the following way. Reading literacy is comprehending, utilizing, assessing, “reflecting on, and engaging with texts” so that one may accomplish their purpose, broaden their horizon, increase their capability, and take part in community (OECD, 2019b, p. 34). Reading is an important skill in different domains of life. For example, it can include areas such as communication, defining locations, participants, time of particular events, or the reasons behind all of these concepts. Therefore, PISA pays close attention to reading literacy since the current age of technology shapes the way people study, work, and communicate. Consequently, it can be observed that the role of digitization is important since the PISA 2018 survey was delivered via computer-based assessment. Thus, digitization underscores the importance of reading as an important skill in student development (OECD, 2019b).

As for science literacy, the OECD defines that it is the capability “to engage with science-related issues” along with the concepts “as a reflective citizen” (OECD, 2016, p. 28). Furthermore, a scientifically educated person is likely to participate in discussions concerning science and technology and is more likely to define phenomena and explain data and confirmation scientifically, as well as assess and create scientific inquiry (OECD, 2019b).

Concerning math literacy, according to the PISA report, it is “students’ capacity to formulate, employ and interpret mathematics” in the course of diverse circumstances(OECD 2019b, p. 27). Math literacy is related to an ability to reason mathematically and apply mathematical concepts, operations, cases, and instruments to interpret and anticipate phenomena (OECD, 2019b). The focus will now be drawn to the history of PISA in Kazakhstan.

1.2 PISA and Kazakhstan

In Kazakhstan, the implementation of an updated curriculum and other changes in the assessment system was done with the intention to develop students' analytical and critical thinking skills (MES, 2016). One of the essential catalysts proposed for introducing a reform was the administration of the PISA 2009 survey and assessments and the careful analysis of associated data. The results from 2009 showed a lack of functional skills among Kazakhstani students. The results from the PISA cycles thereafter revealed the existence of problems in students' analytical skills. For example, Kazakhstan's results in PISA 2012 were lower compared to other countries. Even though there were some improvements in mathematics and science between 2009 and 2012, students' reading performance stayed unchanged across the three-year period. Besides, despite putting efforts into increasing student numeracy, the number of students scoring high in mathematics in 2012 was lower than expected. Furthermore, the PISA 2018 results[1] were worse than PISA 2012 results in all three subjects (Reading, Mathematics, and Science), and if compared with results from 2009, the only improved area was Mathematics, but even this area was lower in the most recent 2018 cycle. Overall, it can be observed that Kazakhstani students' results in PISA across all three cycles demonstrated a general trend of stagnation (OECD, 2019b). Recent research has explored possible reasons for Kazakhstani students' low performance but only in Reading. For example, Muratkyzy (2020)

---

[1] PISA 2015 results involved anomalies such as an oversampling of Nazarbayev Intellectual Schools so were not official (*PISA 2015*. (2016). Informburo.Kz/Stati. https://informburo.kz/stati/kazahstan-ischez-s-mirovoy-karty-urovnya-gramotnosti.html).

explored the student-related and school infrastructure-related factors that contribute to Kazakhstani students’ PISA reading performance. Results suggested that school status as an intellectual school, Russian language of instruction, school socio-economic status (SES), and the percentage of females in the school contributed positively to PISA reading. However, the reasons for these low scores for all subject areas and an examination of other school environmental and contextual factors is yet to be carried out. This suggests that there is generally a dearth of secondary analysis of PISA results that explores the factors affecting student outcomes in Kazakhstan. The current study attempts to fill this gap by exploring the effect of multiple school-climate related factors and their effect on Kazakhstani student and school PISA performance (OECD, 2018).

1.3 Statement of the Problem

The focus of this research is to identify the role that student perceived school climate has on student academic performance in Kazakhstan. To achieve this, students' 2018 PISA survey responses relating to school climate will be utilized in this study alongside student mathematics, reading, and science literacy outcomes.

Previous research has demonstrated that there is a connection between school climate and students' academic performance (Cohen, McCabe, Michelli, & Pickeral, 2009; MacNeil, Prater, & Busch, 2009; Wang & Degol, 2016). To be exact, all three subjects of interest, that is, Reading (Ning, Van Damme, Van Den Noortgate, Yang, & Gielen, 2015); (OECD, 2019a), Mathematics (Kunter, 2013; OECD, 2014), and Science (OECD, 2016), were closely related to school climate. For example, OECD (2019a) states that all school climate factors are important for students’ academic outcomes. However, which of these factors has the largest effect on students’ scores is not yet understood nor explored in Kazakhstan. Furthermore, there is a general lack of secondary analysis of PISA data in the Kazakhstani context so this type of research can be considered as making a unique contribution. To sum, a secondary analysis of the more recent 2018 PISA data provides a useful way to explore the role that school climate might play in improving academic outcomes in Kazakhstan.

1.4 Purpose of the study

The purpose of this quantitative study is to identify the influence of school climate-related factors on students' academic outcomes in Kazakhstan. It aims to identify the influence of school climate-related factors on student academic achievements namely, Math, Reading, and Science results. Importantly, identifying the role that school climate might play on student academic achievement may provide insights into how schools in Kazakhstan might better support student learning.

1.5 Research questions

RQ1. What level of quality of school climate exists for Kazakhstan 15-year-old students?

RQ2. How much do school climate-related factors vary within and between schools in Kazakhstan?

RQ3. What measurement model best describes student experience of school climate in Kazakhstan?

RQ4. What are the school climate-related antecedents of Mathematics, Reading, and Science outcomes in Kazakhstani schools?

1.6 Significance of the study

Despite being a popular topic internationally, research on school climate within the Kazakhstani context is lacking. Little to no research in this area has been conducted on primary or secondary data sources. Consequently, this research is focused not only on identifying the types of school climate-related factors that might drive academic performance but also on providing a general research framework for undertaking a secondary analysis of data in Kazakhstan.

Moreover, the findings of the research might be useful for different stakeholders, including principals, policymakers, and teachers looking to make changes to school policy and school climate with the primary intention of improving the academic outcomes of students.

1.7 Organization of the Study

This first chapter explained the importance of school climate and what is known in the research about its influence on student academic outcomes. The chapter was comprised of the background of the study, statement of the problem, the purpose of the study, research questions, significance of the study, and a brief conceptual framework.

This thesis further covers five chapters. It comprises a literature review that touches upon the school climate factors, a methodology that elaborates on corresponding tools and the ways to measure the influence of school climate, a results section that presents all statistical outcomes, a discussion chapter that explains the meaning of the results, and, finally, a conclusion where all of the points in the thesis are summarized.

2. Literature review

This literature review includes a conceptual framework of the study and a description of the PISA school climate factors, as experienced by students, and recent research on the role that such factors have on student academic outcomes (OECD, 2019a). This chapter comprises four sections. The four sections cover the importance of the school climate, student disruptive behavior, teaching and learning, and the school community.

This literature review makes use of Cooper's (1988) taxonomy as a framework. Table 1 describes all of the categories chosen for all six characteristics of the literature review as embarked upon in this study.

Table 1
Selection of Literature Review Categories Adopted in the Current Literature Review

Characteristic	Categories
Focus	Research outcomes; Research Methods; Theories; Practices or Applications.
Goal	Integration: (a) Generalization, (b) Conflict Resolution, (c) Linguistic Bridge-Building; Criticism; Identification of Central Issues.
Perspective	Neutral Representation; Espousal of Position.
Coverage	Exhaustive; Exhaustive with Selective Citation; Representative; Central or Pivotal.
Organization	Historical; Conceptual; Methodological.
Audience	Specialized Scholars; General Scholars; Practitioners or Policy Makers; General Public.

Note. Specific category chosen for teach characteristic is underlined; table adapted from Cooper (1988).

Each of the chosen categories for the six characteristics of the literature review will now be described in turn. The Focus of this literature review is the relationship between school climate and student outcomes (see Table 1, Research Outcomes), since the research concentrates on identifying the school climate and its influence on student achievement, namely performance in PISA. Therefore, this literature review will attempt to provide a summary of all PISA school climate factors (and corollaries) which have been examined for their influence on student academic performance.

2. As for the Goal, the integration will mainly pertain to a Generalization because this author synthesizes the general studies and findings on school climate.

3. The Perspective is Neutral in this instance, with the author attempting to be as neutral as possible.

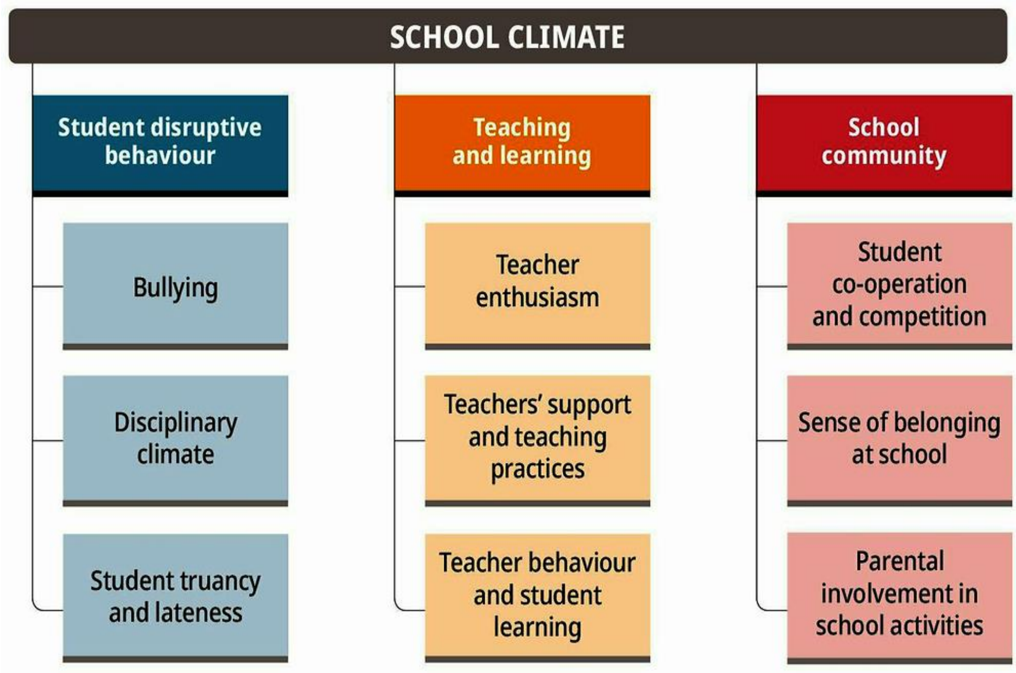
4. The Coverage is concentrated on purposive sampling. The well-cited articles that relate to school climate and student outcomes are first and foremost chosen. Studies that summarize research, such as systematic reviews and meta-analyses will also be a focus of this literature review. This way, central and pivotal articles are selected where possible.

5. The organization of this literature review is constructed conceptually. It covers research that is framed by central theories and concepts on school climate and though generally adapts Figure 1 PISA school climate framework (OECD, 2019a). Furthermore, the school climate framework in OECD (2019a) is relevant given that this study draws upon secondary PISA data. Therefore, Figure 1 was selected as the primary framework for this research.

6. The research is intended for General Scholars, Practitioners, or Policymakers. Concerning General Scholars, insights gained by analyzing the relationship between Kazakhstan school climate and student PISA outcomes might assist them in understand student performance more generally. Regarding the latter audience, Practitioners and Policy Makers, this research may help support schools and school teaching in their planning to create improved conditions that might advance their students' academic outcomes (Randolph, 2009).

Figure 1

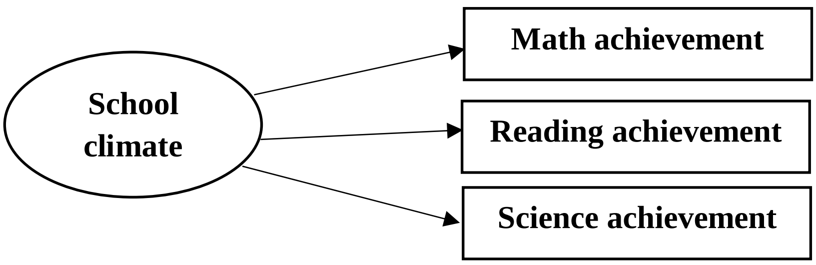
The PISA School Climate Framework



The framework of this research attempts to include all of the school climate related factors, as conceptualized in Figure 2, and measure their influence on Mathematics, Reading, and Science performance of Kazakh students. Therefore, this research follows and adapts a framework from OECD (2019a) which purports the importance of all school climate factors in predicting students' academic scores.

Figure 2

General Model of School Climate on Academic Achievement in PISA



[Independent variables] **[Dependent variables]**

Overall, 12 school climate related factors were used in the PISA survey. However, one set of variables pertaining to "Students' and parents' language competence" ("ST177Q01HA", "ST177Q02HA", "ST177Q03HA") was removed from the study due to not being administered to Kazakhstani students. Therefore, the study included an analysis of 11 school climate related factors and their effect on student academic performance.

2.1 The definition and importance of school climate

According to Cohen et al. (2009) school climate is related to school life, its values, and quality. School atmosphere is built on community's involvement in school activities and indicates "norms, values, interpersonal relationships, teaching and learning practices, and organizational structures" (p. 180). Furthermore, Cohen et al. highlight that school climate is a large notion that incorporates engagement of all stakeholders to have a positive school vision and climate where, every individual is devoted to school activities. Cohen et al. defines that school climate applies to domains of school spirit "(e.g., safety, relationships, teaching and learning, and the environment)" and to broader organizational norms "(e.g., from fragmented to cohesive or "shared" vision, healthy or unhealthy, conscious or unrecognized)" (p. 181).

The definition by Wang and Degol (2016) supports the one provided by Cohen et al. (2009) on the complexity of school environment. Wang and Degol claim that school climate touches nearly every area of school experience and is not only involved with academic improvement but also social interaction, emotional development, and learning independence.

Moreover, research suggests that a school's climate may be a decisive factor that influences the development and change in schools. The findings from a research by MacNeil et al. (2009) support the importance of a positive school climate. In their study, schools with favorable environments regularly demonstrated better student achievement. Therefore, the role of school climate for student academic outcome represents a substantial area of research in

Kazakhstan. As defined in Figure 2, school climate can be conceived as being comprised of three general domains. Each of these three areas will now be described.

2.2 Student Disruptive Behavior

Student disruptive behavior is composed of three general facets: Experience of bullying, the Disciplinary climate, and Truancy and lateness. Each of these facets of student disruptive behavior will now be defined.

2.2.1 Experience of bullying

Safety plays an essential part in every aspect of life, and school education is no exception. Having a safe environment at school is essential to providing a positive school climate and culture. Eccles (2008) argues that students should be taught in a safe environment that does not distort them from studying due to its difficulty and high demand from teachers--a feeling of constant pressure or feeling under threat is not optimal for learning. Therefore, establishing a safe environment at school may be the first step toward academic productivity.

Student safety and student bullying have a close relationship with the leadership practiced at school. According to Kutsyuruba, Klinger, and Hussain (2015), a positive rapport between the principal and school staff can result in healthy relationships between students and a reduction in bullying incidents. Research suggests that the development of confidentiality and the recognition of incidents of cruel behavior should also be established (Kutsyuruba et al., p. 126).

A recent OECD report states that 23% of students mentioned they experienced bullying several times a month, but in Kazakhstan this number was higher and revealed 32% (OECD, 2019c). According to OECD (2019a) results across countries, bullying has one of the most significant effects on students' academic achievements. For example, students who were harassed or threatened at least a few times in a month got 21 points less in reading (OECD, 2019a, p. 46). Moreover, Nansel, Overpeck, Pilla, Ruan, Simons-Morton, and Scheidt, (2001) state students involved in bullying did not only have poorer academic results, but this was also heavily related to a negative school climate caused by its influence on other students. This is because they tend to influence others, and therefore the number of students involved in disciplinary actions increases.

Another important facet of student disruptive behavior is an Anti-bullying attitude. In this matter, students who were frequently bullied might have a weaker Anti-bullying attitude (OECD, 2019a, p. 46). Meanwhile, 88 % of students across OECD countries had a positive Anti-bullying attitude, but less students, 74%, agreed or strongly agreed with helping bullied students (OECD, 2019c).

Therefore, it demonstrates that bullying is a far more complex concept and has bigger consequences than just the deviant behavior of one or a few students. Creating a safe learning environment is a complex task, therefore it requires the active participation of all stakeholders to achieve the desired results. Considering this area is under-researched in Kazakhstan, paying attention to the phenomenon of bullying and the possibility of the importance of anti-bullying attitudes may be useful.

2.2.2 The role of disciplinary climate

Blank and Shavit (2016) conducted research to identify the effect of disruptive academic behavior in the classroom. According to the findings, the authors found that an unpleasant disciplinary climate harmed student outcomes beyond the influence of the control variables, such as student, classroom, school characteristics, as well as students' prior academic performance. In this study, students' test scores in Hebrew were calculated and it was established that students in a comparatively pleasant classroom environment ("one standard deviation below)" the mean interruption level would receive a mean score of 77.7, whereas the students with a higher level of distraction level ("one standard deviation above)" the average would get a score of 73.6 (p. 10). Even though the difference between the scores was not large, it still shows how distraction may affect students' academic outcomes.

Ma and Willms (2004) researched approximately 25,000 Grade 8 students from over 1,000 schools. Analysis suggested that students with a positive classroom climate had higher academic achievements, and in fact, of all the school disciplinary factors, the classroom climate was one of the main predictors of student outcomes. The authors argue when it comes to establishing a favorable disciplinary school climate, the entire school (not the classroom) should be the focus and that such a school environment can be established by the direct involvement of principals and administrators and only then can this transition to teachers and lastly to the classroom level.

Another example of research on the influence of disciplinary climate on student attainment was conducted by Ning et al. (2015). The study was based on PISA 2009 data revealing that 53 out of 63 participant countries demonstrated sufficiently higher scores in reading, whereas in the other 12 countries the results were insignificant. The authors make the case that there was a moderating effect where the absence of positive academic outcomes during good classroom conditions was explained by large class sizes. In this instance, the authors make the case that teachers seem to spend more time establishing the class climate but lose valuable time for teaching and organizing class materials.

As for Kazakhstan, results suggest that teachers do not have to waste much time to calm their students--only 8% of students reported this as an issue. It should be noted that classroom noise in Kazakhstan is more than thrice less than in the OECD average (26%). After controlling for SES ("ESCS") which stands for "Index of economic, social and cultural status", results suggested that students that experienced regular noise in every lesson (or in most lessons) scored 26 fewer scale score points in reading (OECD, 2019c).

2.2.3 The role of student truancy and lateness

As for truancy in Kazakhstan, students were more than twofold prone to skip one school day (52%) than the OECD average (21%). However, concerning lateness, the OECD and Kazakhstan averages were 48% and 52%, respectively. In the majority of OECD countries, poorer truancy was noticed among students experiencing bullying. On the other hand, students with a better sense of school belongingness and those receiving support from parents tended to attend the class more (OECD, 2019c).

According to research, targeting absenteeism is more important than extending the academic calendar. For example, a ten-day increase in the school calendar may only improve students' numeracy and literacy skills by 1.7% (0.8% of a standard deviation) while reducing truancy may increase these numbers to 5.5% (2.9% of a standard deviation), correspondingly. Furthermore, research suggests that the detrimental effect of absenteeism varies with age. For example, one-day truancy in 5th grade has three times the negative effect for math performance than it does in 3rd grade. Lastly, tackling students' skipping of classes may help decrease the gap separating high and low student academic performance (Aucejo & Romano, 2016).

In Ohio, a study was done on the relationship between student attendance and student outcomes, which included students from fourth, sixth, ninth, and twelfth grades. In all grades, at least a moderate relationship was found with the strongest relationship exhibited among ninth-grade students (Roby, 2004). Considering the participants of PISA are mostly ninth-grade students, paying attention to attendance seems to be important.

2.2.4 Summary

In summary, research suggests that student disruptive behavior has a significant effect on student academic performance. Therefore, close attention should be paid to improve school safety, considering, there are more bullying cases in Kazakhstan than in the OECD average (OECD, 2019c). Furthermore, it was established that improving classroom climate was important in increasing student academic scores, but this was not enough to build a positive school climate (Ma & Willms, 2004). School attendance was found essential, and its significance increases with the student progressing through grades (Aucejo & Romano, 2016). It can be concluded that, students who do not experience bullying and distraction, and keep good tracking are likely to be more successful with their studies.

2.3 Teaching and Learning Environment

Teacher enthusiasm, support, and verbal and non-verbal teacher behavior constitute an important part of the class and school climate. Each of these facets of school climate will now be discussed in turn.

2.3.1 The role of teacher enthusiasm

Teacher motivation also plays an all-important role in school climate. In this regard, teacher job satisfaction is important to provide a positive environment for all students (Frase & Sorenson, 1992).

There is a positive connection between teacher and student motivation. In fact, teacher motivation can be considered as a driver of student gratification (Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009; Kalyar, Ahmad, & Kalyar, 2018). Furthermore, research has found that motivated teachers were better at dealing with diverse school tasks including diverse subject matter and improving student academic performance (MacNeil et al., 2009; Kalyar et al., 2018). Kunter (2013) defines teacher enthusiasm as constituting two categories: enthusiasm for a subject and enthusiasm for teaching. According to Kunter’s research, devotion for the subject was not as important as enthusiasm for teaching. Therefore, enjoyment of interactions with students could be stated as being more significant than mere love for a subject.

Building a positive school climate might be arduous and this might be especially difficult if a toxic culture has been practiced for many years. To solve this issue, schools need leaders, change agents, and perseverance, and teachers play a vital role in initiating such transformations. To this point, Deal and Peterson (1998) identified some examples and common features that schools with positive school cultures have. The researchers give examples of schools that developed and even redesigned their schools with a vision toward an improved climate. For example, Ganado Primary School in Ganado, Arizona, was initially one of the schools with a poor school climate and culture. However, the school now is a place with a positive environment. Driven by the impact of a new principal and leadership, an improved school climate resulted in an increase in student motivation for academic achievement (Deal & Peterson).

To note, an improved school climate is not always associated with school socio-economic status and affluence. Deal and Peterson (1998) emphasize that it is important to celebrate small successes and to stimulate teacher motivation as it may occur as one of the significant drivers of high student accomplishment.

2.3.2 The role of teachers' support and teaching practices

Federici and Skaalvik (2013) defined emotional and instrumental support as two dimensions of teacher support. In this matter, Emotional support is associated with compassion, kindness, incentive, respect, and nurture, while instrumental support represents "tangible support", such as, teachers assisting students in addressing a task or attaining a challenging issue (p. 21). The researchers found that emotional and instructional support are strongly correlated. It means teachers tend to pay attention to students' well-being, as well as to their improvement in the subject.

According to results, teacher support positively affects student commitment. In a favorable environment with a caring teacher, students are more prone to attend lessons and perform better in tests. For instance, secondary school students who reported low teacher support levels were 68% more predisposed to be withdrawn from school (Klem, & Connell, 2004).

Another essential factor for establishing a pleasant school environment is the building of interpersonal relationships among diverse stakeholders. A positive school climate provides favorable opportunities for students. Consequently, the school can (1) support students by acknowledging and appreciating their abilities, as well as (2) establish more friendly relationships between students and teachers. Concerning this matter, the role of teachers and administrators should be defined as they are the main agents for creating positive relationships at school (MacNeil et al., 2009).

2.3.3 Teacher/student behavior and student learning

The types of teacher behavior have connections with the types of skills students acquire. Research on the effect of teacher behavior on student learning utilized pre- and post-tests to support this outcome. According to the findings, learning of facts was strongly related to the teacher being clear, expressive, and good at lecturing; and, gains in comprehension were linked substantially with the teacher being energized and flamboyant (Solomon, Rosenberg, & Bezdek, 1964).

Different factors influence teacher behavior. Several studies have shown the connection between teacher and student behavior (Klein, 1971; Sherman & Cormier, 1974). According to Klein (1971), student behavior is one of the more important corollaries of the way teachers act in the classroom. It was established that teacher behavior tends to be associated with verbal and non-verbal student behavior.

The research has also demonstrated that teachers tend to behave differently toward incidents of student failure. Georgiou, Christou, Stavrinides, and Panaoura (2002) used structural equation modeling to identify that teachers show more compassion and less annoyance when they relate student's inferior progress to their low competence, while they tend to behave more aggressively when linking low student achievement to students' low level of efforts. This finding is important because it explains the reason behind teachers' aggression and compassion. It also suggests that low-achieving students are not left behind, and teachers are always there to support them.

In a study by Kyriakides, Creemers, and Antoniou (2009), the role of teacher behavior on student achievement was described in a detailed way. The authors identified five types of teacher behavior. In this regard, the first type is related with direct teaching, the second type incorporates both the direct and active teaching, the third type deals with implementation of active teaching, the fourth type incorporates using differentiation method, while the fifth type comprises accomplishing a quality and differentiation by utilizing diverse methods. Findings suggested that student academic outcomes improved as teacher behavior become more advanced (pp. 19-20). Therefore, teacher behavior appeared to have a direct influence on student learning.

2.3.4 Summary

In conclusion, it can be uttered that teacher devotion for the subject was found to be not as important as devotion for teaching. In this regard, the role of principals is essential in building a positive school climate among teachers and students. Concerning the teacher support, it was not only important for students' studies but also essential for students' commitment to the school. The role of teacher behavior could vary depending on students' abilities, and there are more types of teacher behavior that need to be further researched.

2.4 School community

The School community is conceived as constituting aspects of student cooperation, a sense of belonging, and parental involvement in school activities. A summary of the research concerning how these aspects of the school community affect student learning outcomes will now be provided.

2.4.1 Student co-operation and competition

Collaboration is essential both for teachers and students. While cooperating, teachers increase their involvement in lessons. According to Deal (1998), cooperative teachers are more likely to feel responsible for students' academic outcomes. Thus, this illustrates the essence of cooperation.

Integrating student cooperation and competition serves another important part in establishing a school's climate and culture. According to Tauer and Harackiewicz, (2004), combining competition and cooperation has a consistent effect on better task performance levels and motivation. Meta-analytic research on the roles of competition and cooperation on student achievement suggests that cooperation is more essential than competition concerning student achievement and productivity. The positive effect was across all subjects and cooperation was effective when it lasted for a short period. The only point when cooperation was ineffective was related to rote decoding and decoding tasks (Johnson, Maruyama, Johnson, Nelson, & Skon, 1981). Thus, this illustrates the essence of cooperation and its diverse and broad nature and relationship with learning outcomes.

Agasisti (2011) could not find a clear effect of competition on Italian students' academic outcomes. According to the available data, even the broadly piloted region in Italy (the Lombardy Region) depicted only a small to negligible effect. Therefore, this suggests that the promotion of competition in schools is more complex and requires further exploration in additional educational contexts.

2.4.2 Sense of belonging at school

Students' sense of belonging at school has been associated with a diverse set of factors including students' academic performance. Based on the students' responses to questions about school satisfaction, students were divided into three categories: low, middle, and high school satisfaction groups. Results of the subsequent analysis suggested that students' academic achievements were strongly associated with their level of fulfillment of school belonging (Huebner & Gilman, 2006).

Research based on PISA 2003 data in Mexico looked at the importance of student school engagement on academic achievement. The study found a strong relationship between these two factors. On the other hand, counterintuitively, students' strong commitment to the teacher did not enhance their performance--in fact, while controlling for several other covariate effects, it tended to be related to lower student outcomes (Weiss & García, 2015). The results of this study reveal that further research into this topic is warranted. Nevertheless, students' sense of belonging appears to have important social implications in schools and maybe indirectly associated with academic outcomes and student wellbeing.

Finally, a meta-analytic study by Allen, Kern, Vella-Brodrick, Hattie, and Waters (2016) reviewed multiple studies in order to find the strongest predictors of a sense of belonging at school. Regarding the study, it was discovered teachers' support and teachers' positive personal characteristics were the main drivers. Interestingly, these impacts were more evident in rural areas than in urban regions.

2.4.3 Parental involvement in school activities

The role of parents in school life should also be paid due attention. Hampden-Thompson, Guzman, and Lippman (2013) analyzed PISA results from 21 countries. The authors identified the importance of parental school involvement for student reading literacy. However, specifically, results suggested that parental help with student homework was insufficient for improving student outcomes. In fact, research by the authors, Hampden-Thompson et.al (2013), demonstrated that highly frequent parental involvement in students' homework tasks substantially reduces their literacy scores. The authors suggest that the best approach that could improve students' success is believed to be balanced assistance from parents (also see, Benner et al, 2016). However, more research was necessary to identify the specific parental behaviors that might function to best support student learning.

To this end, further meta-analytic research by Castro, Expósito-Casas, López-Martín, Lizasoain, Navarro-Asencio, and Gaviria (2015) suggested the following as the substantial predictors of student academic achievements: parents' high expectations, communication about school-related activities, and practicing reading skills. These specific forms of parental involvement in education proved to develop students' success at school, while parent volunteering did not afford any substantive improvements (Wang & Sheikh-Khalil, 2013).

2.5 Conclusion

The summary of the prominent studies in the field suggests that multiple school climate related factors are significant predictors of student academic outcomes. However, it was established that some of the factors were not as important as it was believed. For example, the role of competition was not as significant as cooperation. Further, parental involvement with student homework seemed to cause detrimental effects for students' academic outcomes if there was excessive support. In addition, it was identified that dealing with student truancy and targeting absenteeism is more essential than extending the academic calendar. To sum up, multiple school climate related factors, except for student competition and specific forms of parental involvement, tended to result in positive social outcomes and tended to be important drivers of student academic achievement. Given the dearth of research on the topic in Kazakhstan, the following research questions are specified for the current study:

RQ1. What level of quality of school climate exists for Kazakhstani 15-year-old students?

RQ2. How much do student-level school climate-related factors vary within and between schools in Kazakhstan?

RQ3. What measurement model best describes student experience of school climate in Kazakhstan?

RQ4. What school climate-related factors are associated with Mathematics, Reading, and Science outcomes in Kazakhstani schools?

The following chapter provides an explanation of the methodological approach used to answer the four research questions above.

3. Methodology

3.1 Introduction

This chapter provides an explanation of the methodology adopted to answer the research questions in the current study. The chapter provides an explanation of research design and rationale, sampling, data analyses, methods, ethics, and limitations. The purpose of this quantitative study is to identify the influence of school climate-related factors on students' academic outcomes in Kazakhstan.

Research Questions

This research covers four questions:

RQ1. What level of quality of school climate exists for Kazakhstani 15-year-old students?

RQ2. How much do student-level school climate-related factors vary within and between schools in Kazakhstan?

RQ3. What measurement model best describes student experience of school climate in Kazakhstan?

RQ4 What school climate-related factors are associated with Mathematics, Reading, and Science outcomes in Kazakhstani schools?

The methodology section that follows presents details pertaining to the research design, sampling, instruments, data analysis, and ethical issues specific to the current study.

3.2 Research Design

According to Creswell (2014), "a non-experimental research design is applied to observe and identify the relationship between two or more related variables as carrying out controlled experiments is not often considered feasible" (p. 364). In the current study, there are multiple independent variables (both student- and school-level variables) and three dependent variables of interest, i.e., the student outcomes of Math, Reading, Science. The independent variables in this study capture school climate and they are all taken from the 2018 PISA student and school surveys. The dependent variables in this study include Mathematics, Reading, and Science scale scores from the publicly available PISA 2018 student datasets. Student-level variables are defined as those that vary for each student, for example, student socio-economic status, whereas school-level variables are defined as those that vary for each respective student in that school, for example, the extent to which a school reports to implement certain equity-related policies. Therefore, because of the nestedness of data in this study, multilevel modelling was employed as the primary method of analysis (Bates, Maechler, Bolker, & Walker, 2015). However, in the case that there would not be sufficient between-school variation in school climate, the appropriate method would default to a single-level form of multivariate analysis.

3.2.1 The Rationale for Choosing PISA Data

The transition between middle and high school has been widely recognized as a critical phase in the development of younger learners (Eccles, 2008). Therefore, the analysis of 15-year-old Kazakhstani student performance in PISA may provide valuable developmental insights. For this reason, it is meaningful to look at the role of contextual factors and patterns of achievement of students who partook in the PISA examinations in Kazakhstan.

From 2007, on account of educational reforms, Kazakhstan started taking part in international tests, one of which was PISA (IAC, 2017, p. 160), which aimed to assess the quality of education and provide an objective and independent judgment of the strengths and weaknesses of school education in the region. The intention was that Kazakhstan would use the results of PISA to identify the main areas for educational development. For example, “the below-average achievements in PISA of Kazakhstani students have generated pressure for reform and have led the educational experts in Kazakhstan to be sent to visit other countries to learn how to refine its educational policies” (Yakavets, 2014, p. 45). The results of these large-scale assessments revealed an unacceptably low level of knowledge and abilities (OECD, 2018), and it was argued that this would lead to low demand in the world labor market

(Tasbulatova, Belosludtseva, & Grooting, 2007, p. 8). Therefore, in response to the reports and analysis of the results, policymakers initiated several reforms intended to modernize the Kazakhstani curriculum.

The current study is focused on PISA results, as this test is one of the main tools for assessing Kazakhstani students' functional and critical thinking skills. As stated, early results from PISA have guided more recent curricula reform. Specifically, more recent results have suggested that despite Kazakhstani students demonstrating improved results in mathematics and science, they still had difficulties with tasks that required critical thinking skills (OECD, 2018). It was speculated that this was the result of routine learning processes in Kazakhstani schools that concentrated mainly on memorization and did not involve enhancements in functional literacy (OECD, 2018).

In conclusion, secondary analysis of recent Kazakhstani student PISA results may provide useful observation of the role that school climate might have on student academic performance. Insights into the role that school ambience has on students' academic performance in PISA may have important implications for understanding academic learning in general in the country enabling schools to set new policy and practice based on empirical evidence.

3.3 Sampling and Data Preparation

Since the research was based on secondary data, there was no need for this author to collect data. The data was obtained from the OECD official website ("PISA 2018 Database", n.d.) and used a multi-stage cluster sampling methodology. However, certain data preparatory steps were undertaken using the PISA 2018 data which was downloaded in SPSS .sav format.

3.3.1 Data Extraction and Preparation

This author made use of the open-source R statistical software for the data preparation and analysis of the PISA data. In R (R Core Team, 2019), the SPSS files in .sav format were read in with the assistance of the haven package (Wickham & Miller, 2020). The participants of the survey were the 19,507 anonymized 15-year-old school students from 616 schools in Kazakhstan. Kazakhstani school and student data were extracted and merged by “CNTSCHID” which included the common prefix number of “398” for Kazakhstan. After reading in both student and school datasets, data were merged. This data included 19,507 students from 616 schools inclusive of 1313 variables. Since the “ESCS” was considered to be an important antecedent of student performance, this variable was selected for analyses as well as gender. However, ESCS was only an individual level variable and not associated with each school. Therefore, there was a need to create average SES for each school and this was defined as “AVG.PISA.ESCS”.

After analyzing all school climate student-level variables, 904 students and 190 schools were removed due to missing data and singleton answers (i.e., incidents where there was only one student respondent from a school) reducing the sample to 11,528 students in 426 schools. In order to account for possible anomalous results from very small numbers of students in schools, schools with fewer than 10 students were also removed (Lai & Kwok, 2014). Therefore, the final data for conducting the *main multi-level analyses* included 399 schools and 11,317 students. For the current analysis, only the variables of interest concerning school climate as well as some common school-level variables were selected. Overall, 70 variables were chosen. All the semantically negative questions were reverse coded. All the negatively worded questions are earmarked with an “R” in Table 2. Then all of the means for each variable will have the same meaning.

Table 2

Student-Level Independent Variables from the PISA Dataset [Instrument 1]

		(1) Experience of bullying[1]
1	ST038Q04NA	Other students made fun of me.
2	ST038Q05NA	I was threatened by other students.
3	ST038Q06NA	Other students took away or destroyed things that belonged to me.
4	ST038Q07NA	I got hit or pushed around by other students.
		(2) Anti-bullying attitude[2]
5	ST207Q01HA	It irritates me when nobody defends bullied students.
6	ST207Q02HA	It is a good thing to help students who can't defend themselves.
7	ST207Q03HA	It is a wrong thing to join in bullying.
8	ST207Q04HA	I feel bad seeing other students bullied.
9	ST207Q05HA	I like it when someone stands up for other students who are being bullied.
		(3) [R] Disruptive Student Academic Behavior[3]
10	ST097Q01TA	Students don't listen to what the teacher says.
11	ST097Q02TA	There is noise and disorder.
12	ST097Q03TA	The teacher has to wait a long time for students to quiet down.
13	ST097Q04TA	Students cannot work well.
14	ST097Q05TA	Students don't start working for a long time after the lesson begins.
		(4) Student truancy and lateness[4]
15	ST062Q01TA	I <skipped> a whole school day.
16	ST062Q02TA	I <skipped> some classes.
17	ST062Q03TA	I arrived late for school.
		(5) Teacher enthusiasm[2]
18	ST213Q01HA	It was clear to me that the teacher liked teaching us.
19	ST213Q02HA	The enthusiasm of the teacher inspired me.
20	ST213Q03HA	It was clear that the teacher likes to deal with the topic of the lesson.
21	ST213Q04HA	The teacher showed enjoyment in teaching.
		(6) [R] Teacher's support and teaching practices[3]
22	ST100Q01TA	The teacher shows an interest in every student's learning.
23	ST100Q02TA	The teacher gives extra help when students need it.
24	ST100Q03TA	The teacher helps students with their learning.
25	ST100Q04TA	The teacher continues teaching until the students understand.
		(7) Teacher behavior and student learning[2]
26	ST211Q01HA	The teacher made me feel confident in my ability to do well in the course.
27	ST211Q02HA	The teacher listened to my view on how to do things.
28	ST211Q03HA	I felt that my teacher understood me.
		(8) Student co-operation[5]
29	ST206Q01HA	Students seem to value cooperation.
30	ST206Q03HA	Students seem to share the feeling that cooperating with each other is important.
31	ST206Q04HA	Students feel that they are encouraged to cooperate with others.
		(9) Student competition[5]
32	ST205Q01HA	Students seem to value competition.
33	ST205Q02HA	It seems that students are competing with each other.
34	ST205Q03HA	Students seem to share the feeling that competing with each other is important.
35	ST205Q04HA	Students feel that they are being compared with others.
		(10) [R]Lack of sense of belonging at school[2]
36	ST034Q01TA	I feel like an outsider (or left out of things) at school.
37	ST034Q04TA	I feel awkward and out of place in my school.
38	ST034Q06TA	I feel lonely at school.
		(11) Parental involvement in school activities[2]
39	ST123Q02NA	My parents support my educational efforts and achievements.
40	ST123Q03NA	My parents support me when I am facing difficulties at school.
41	ST123Q04NA	My parents encourage me to be confident.

Note. [1]Denotes scales using [1]1= Never or almost never, [1]2= A few times a year, [1]3= A few times a month, [1]4= Once a week or more,[2]1= Strongly disagree, [2]2= Disagree [2]3= Agree, [2]4= Strongly agree; [3]1 = Never or hardly ever, [3]2 = Some lessons, [3]3 = Most lessons, [3]4 = Every lesson; [4]1=Never, [4]2=One/two times, [4]3=Three/four times, [4]4=Five/more times; [5]1= Not at all true, [5]2= Slightly true, [5]3= Very true, [5]4= Extremely true.
[R] The entire presented scales were reverse-coded to have appositive meaning. The original scales were ([R] Lack of Disruptive Student Academic Behavior[3], [R] Lack of Teacher's support and teaching practices[3], [R]Sense of belonging at school[2]). After reverse coding all scales have the codes from positive to negative.

3.4 Data collection tools

As stated, two questionnaires are utilized to conduct this research. The student and school questionnaires, available from the official OECD website, are utilized for data collection (OECD, 2018). The dependent variables in the study are the first five plausible values for each student for each of the three subjects of interest in this study, namely Mathematics, Reading, and Science ability. The student independent variables are now presented.

3.4 1 Instrument: Student Questionnaire

The student-level independent variables included in this study (completed by individual Kazakhstani students) are presented in Table 2.

The student questionnaire covers 41 school climate questions specific to scales about students' (1) Experience of Bullying, (2) Anti-Bullying Attitude, (3) Disruptive Student Academic Behavior, (4) Student Truancy and Lateness, (5) Teacher Enthusiasm, (6) Teacher's Support and Teaching Practices, (7) Teacher Behavior and Student Learning, (8) Student Co-operation and (9) Competition, (10) Sense of Belonging at School, and (11) Parental involvement in school activities. As noted in Table 2, Scale/Response Options include both frequency and agreement formats.

3.5 Data analysis

3.5.1 Initial Assessment of Data

Initial data preparation included an assessment of data completeness. After confirming this, an examination of the skewness of each variable followed. Where necessary, normality transformations involved the application of the optimal exponent (Box & Cox, 1964).

3.5.2 Descriptive Analysis

RQ1 asks What level of quality of school climate exists for Kazakhstan 15-year-old students? To answer this question, descriptive statistics were provided for all variables included in this study. In fact, descriptive statistics included 426 schools and 11,528 students. This involved presentation of min, max, mean, standard deviations, skewness for each variable to be modeled in the current study. Descriptive statistics were run using base R functions (R Core Team, 2019), which did not require the installation of additional packages. However, to check for variable skewness, the e1071 package was used (Meyer, Dimitriadou, Hornik, Weingessel, & Leisch, 2020). Descriptive statistics for student gender ("ST004D01T") and SES ("ESCS") were also provided. The gender was coded as 1 = female and 2 = male. Furthermore, descriptive statistics included the mean of the first five plausible values (PVMMATH, PVMSCIE, PVMREAD) for all PISA scores (Mathematics, Reading, and Science).

3.5.3 Application of Multi-level Modelling

RQ2 asks, How much do student-level school climate-related factors vary within and between schools in Kazakhstan? To answer this question, the lme4 package was used for estimating the ICCs (intra-class correlations) for each of the independent variables (Bates, Maechler, Bolker, & Walker, 2015), and the optimx package for supporting lme4 package's functions (Nash & Varadhan, 2011; Nash, 2014). Design effects (*de*) were also estimated to determine the necessity of multilevel modelling. Design effects were calculated in accordance with $de = 1 + ICC(c-1)$, where ICC = intra-class correlation, c = average number of students sampled per cluster are also provided at each time point. ICC values over .10 and design effects over 2.0 provide evidence for substantive effects due to school clusters (Goldstein, 2003).

3.5.4 Confirmatory Factor Analysis

RQ3 asks, What measurement model best describes student experience of school climate in Kazakhstan? To answer this question, this study makes use of confirmatory factor analysis. Confirmatory factor analysis (CFA) and structural equation modelling (SEM) were adopted as a general technique in this study to assess the validity of the substantive theory in the field school climate and student achievement. A two-step approach involving the specification of separate measurement and structural models was undertaken (Anderson & Gerbing, 1988).

Both CFA and SEM were undertaken with the assistance of the R lavaan package (Rosseel, 2012).

Minimum standardized item-factor loadings were set at b = .40 (Ab Hamid, Sami, & Mohmad Sidek, 2017). Inter-factor correlations (r) were also interpreted alongside the shared variance (r^2). Further, the average variance extracted (AVE) was calculated in accordance with

$$AVE = \frac{\sum_{i=1}^{k} \lambda_i^2}{\sum_{i=1}^{k} \lambda_i^2 + \sum_{i=1}^{k} Var(e_i)} \quad [1]$$

where, k is the number of items, λ^2 is the item-factor loading of item i and $Var(e_i)$ is the variance of the error of item i, where,

$$Var(e_i) = 1 - (\lambda_i^2) \quad [2]$$

To demonstrate convergent validity, AVE values for each construct should generally exceed .50 (Ab Hamid et al., 2017) and the related inter-factor variance should be less than the average variance extracted (r^2) < AVE. An inter-factor correlation matrix was also presented for the factors in the measurement model. The dependent variables, namely, PVMMATH, PVMSCIE, and PVMREAD, were also included so as to interpret the bivariate relationships between these variables.

The CR reliability coefficient is used to determine the reliability of a single scale. The composite reliability (ρ_c) of a single scale can be calculated as follows,

$$\rho_c = \frac{(\sum_{i=1}^{k} \lambda_i)^2}{(\sum_{i=1}^{k} \lambda_i)^2 + \sum_{i=1}^{k} Var(e_i)} \quad [3]$$

To demonstrate convergent validity the related inter-factor variance should be less than the AVE, and the CR should be above .70 (Ab Hamid et al., 2017).

From within a confirmatory framework, researchers are well-advised to report and consider multiple-model fit indices including the χ2/df ratio (under 3.83) and associated non statistically significant p value Walker (2013), SRMR (below .08) (Hu & Bentler, 1999), RMSEA (below .08) Browne and Cudeck (1989, 1992) and Byrne (2001), CFI (above .90) and

TLI (above .90) (Hoyle, 1995)., and gamma hat (above .90) when considering model fit (Byrne, 2001). In addition, based on their popularity, the WRMR (close to 1.00) fit statistic (Yu & Muthén, 2002) and Cronbach's alpha values (above .70) for factors' construct validity are also reported throughout this study.

To address RQ4 concerning, as part of the second step in the aforementioned two-step approach, this study asks, What school climate-related factors are associated with Mathematics, Reading, and Science outcomes in Kazakhstani schools, this study employs structural equation modelling (SEM). At this stage, all three academic outcomes are modelled as dependent variables while all of the independent variables relating to school climate are modelled as independent variables. In this way, this research identifies the most prominent school-climate related antecedents for the three student academic outcomes. A rule of ($b < .10$) which means if the beta weight is .10, then this only explains 1% of the variance in the outcome. Categories, which are below the beta minimum value ($b < .10$) can not be considered as driver factors. Finally, at this stage, model fit indices will also be inspected for the structural model in order to assess the fit of the data to the model. To note, the entire R script for all data preparation and analysis steps is provided in the Appendix.

3.6 Ethical Issues

All research in this study was undertaken in accordance with the NUGSE code of ethics. The research did not include the financial cost to participate. The research does not analyze individual-specific student or school responses but rather builds general results.

Since this research uses secondary data sources, there is an extremely low risk to its participants. Further, it did not pose difficulties to participants and did not include vulnerable human subjects (Ispambetova, 2018). Finally, the results identify general patterns in the data with no individual school or person being identified.

3.7 Limitations

First, there was a very limited amount of supporting literature in the Kazakhstani context since this research area is not common among researchers. Therefore, this study has very little literature to guide it. In addition, due to the nature of the study involving secondary data, there was an inability to change or adapt some of the questionnaires. Due to the above-mentioned points, there were some restrictions with this research. For future studies, it is recommended that researchers conduct investigations on primary source data to further advance the field.

3.8 Conclusion

This methodology chapter provides a summary of the sampling used in PISA, the use of secondary data, the data preparatory procedures, and the statistical procedures used to answer the four research questions, namely, descriptive statistics, multilevel modelling, confirmatory CFA, and SEM. This study did not include direct involvement with human subjects since it was based on secondary data.

4. Results

4.1 Introduction

This chapter presents the findings of the research. The purpose of the study is to identify the role of school climate on student academic achievement based on the following four questions:

RQ1. What level of quality of school climate exists for Kazakhstan 15-year-old students?

RQ2. How much do student-level school climate-related factors vary within and between schools in Kazakhstan?

RQ3. What measurement model best describes student experience of school climate in Kazakhstan?

RQ4. What school climate-related factors are associated with Mathematics, Reading, and Science outcomes in Kazakhstani schools?

The research questions were answered by utilizing PISA OECD 2018 dataset. All the findings are described in detail in the following sections.

4.2 Analysis steps

The Student Questionnaire from PISA 2018 (OECD, 2019) was synthesized to identify the school climate factors. After data preparation, descriptive statistics were carried out on 11,528 students. However, due to the possibility of anomalous results from small school samples, schools with less than ten students were removed resulting in 11,317 students. This data was used for the confirmatory factor analysis (CFA) to measure the validity of students' responses on school climate. The main research questions, RQ3 and RQ4, were answered by using CFA and full SEM models which ultimately validated and identified the main drivers of student performance in PISA mathematics, science, and reading in Kazakhstan.

4.3 Results

This chapter presents the results for each of the four research questions:

RQ1: A moderate level of school climate quality exists for Kazakhstani 15-year-old students.

RQ2: Variance in student perceived school climate varies largely within schools but not between schools in Kazakhstan.

RQ3: A ten-factor measurement model provides a useful representation of student-perceived school climate factors in Kazakhstan.

RQ4: The majority of school climate factors were found to contribute significantly to all three academic outcomes, but the most important drivers were students' Experience of bullying ("ST038"), which had a substantive negative effect, and Anti-bullying attitude ("ST207"), which had a substantive positive effect.

4.4 RQ1: Prominent student school climate factors in Kazakhstan

Descriptive statistics were used to answer this question and analyze the data (Thompson, 2006). All the student-level school climate items are presented in Table 3. The student questionnaire overall covers 41 school climate questions concerning school climate in addition to descriptive statistics for gender and ESCS as demographic variables of interest and PVMMATH, PVMSCIE, PVMREAD as outcome variables.

According to Table 3, it can be observed that all class climate factors have equal min and max values. In this matter min = 1 and max = 4. It is noted that *M*(mean) for all items pertaining to factors of a negative nature range from (*M=1*.33 to *M*=2.03). For example, the means for student Experience of bullying (ST038) has a mean from (*M*=1.33 and *M*=1.45). In this regard, overall about 8% of students reported being bullied once a month, however, out of all bullied students 33% were bullied at least once a month, whereas in OECD it equaled 23% (OECD, 2019a). Disruptive Student Academic Behavior (ST097) had a mean from (*M*=1.45 to *M*=1.68). For example, disobeying the teacher and noise in the classroom was reported by 11% and 9% of students respectively, whereas the number for OECD average was about three times higher (OECD, 2019a).

Mean scores for Student truancy and lateness (ST062) ranged from (*M*=1.68 to *M*=1.74), where about 49 % of students missed school at least one or two times, whereas the OECD average was less than twice equaling to 21 % (OECD, 2019a). Concerning being late to school about 51 % of students reported this issue, with OECD data being quite close at 48%. Lack of sense of belonging at school (ST034) had a mean range from (*M*=1.97 to *M*=2.03), are all low suggestive of low levels of such phenomenon. For example, one out of five students feels a lack of sense of belonging, which results in about 80% of students feel a sense of belonging to school with the OECD average at 75% (OECD). All the remaining class climate factors with a positive nature have a mean from (*M*=2.47 to *M*=3.45). In this matter, Teacher enthusiasm (ST213), Teacher's support and teaching practices (ST100), Parental involvement in school activities (ST123) have more positive responses than others with an average mean of above (*M*=3.00), whereas student competition (ST205) has the lowest mean of all conceivably positive factors at around (*M*=2.47-2.55). The mean for anti-bullying attitude is quite higher than student competition ranging from (*M*=2.75-2.93). As for the standard deviations for all variables, there were no outliers and all the class climate factors did not exceed 1.00 with *SDs* ranging from (*SD*=0.7 - 0.9). Regarding skewness, the only Experience of bullying (ST038) could be considered substantively skewed with the value just above |2.00|.

Compared to the OECD average, these results suggest that the school climate in Kazakhstan is not much different when it comes to a Lack of sense of belonging at school (ST034), Teacher enthusiasm (ST213), Teacher's support and teaching practices (ST100). In Kazakhstan, about 87% of students reported getting support from parents, whereas the OECD average was about 70%. A more positive school climate in Kazakhstan exists concerning Disruptive Student Academic Behavior (ST097). Despite the lateness not causing many issues truancy seemed to be a problematic area in Kazakhstan.

Table 3

Descriptive Statistics for Student-Level Variables

Variable ID	description	min	max	Mean	SD	Skew	ICC	D Eff	Codes
ST004D01T	Gender	1	2	1.49	0.50	0.03	0.044	2.15	1,2
ST097Q01TA	**A.** Frequency of not listening to what teacher says	1	4	1.68	0.79	1.18	0.048	2.25	3
ST097Q02TA[R]	**B.** Frequency of noise and disorder	1	4	1.59	0.73	1.26	0.055	2.43	3
ST097Q03TA[R]	**C**. Frequency of teach waiting for students to quiet down	1	4	1.49	0.75	1.63	0.047	2.22	3
ST097Q04TA[R]	**D**. Frequency of students not working well	1	4	1.49	0.71	1.53	0.031	1.81	3
ST097Q05TA[R]	**E**. Frequency of students not starting work after the les begins	1	4	1.45	0.71	1.73	0.040	2.04	3
ST100Q01TA[R]	**A.** Frequency of teacher showing an interest in every student's learning	1	4	3.19	0.89	-0.81	0.044	2.15	3
ST100Q02TA[R]	**B.** Frequency of the teacher giving extra help when students need it	1	4	3.32	0.84	-1.00	0.026	1.68	3
ST100Q03TA[R]	**C.** Frequency of the teacher helping students with their learning	1	4	3.45	0.78	-1.34	0.035	1.91	3
ST100Q04TA[R]	**D.** Frequency of the continuing teaching until the students understands	1	4	3.37	0.83	-1.14	0.027	1.70	3
ST211Q01HA	**A.** The teacher made me feel confident in my ability to do well in t	1	4	2.88	0.92	-0.77	0.032	1.83	2
ST211Q02HA	**B.** The teacher listened to my view on how to do things	1	4	2.82	0.90	-0.66	0.029	1.76	2
ST211Q03HA	**C.** I felt that my teacher understood me.	1	4	2.89	0.92	-0.75	0.031	1.81	2
ST213Q01HA	**A.** It was clear to me that the teacher liked teaching us.	1	4	3.01	0.81	-0.84	0.058	2.51	2
ST213Q02HA	**B.** The enthusiasm of the teacher inspired me.	1	4	2.93	0.83	-0.64	0.056	2.46	2
ST213Q03HA	**C.** It was clear that the teacher likes to deal with the top of the les.	1	4	3.04	0.76	-0.91	0.028	1.73	2
ST213Q04HA	**D.** The teacher showed enjoyment in teaching.	1	4	3.02	0.81	-0.82	0.032	1.83	2
ST034Q01TA[R]	**A.** I feel like an outsider (or left out of things) at sch.	1	4	1.97	0.85	0.74	0.008	1.21	2
ST034Q04TA[R]	**B.** I feel awkward and out of place in my school	1	4	2.03	0.80	0.67	0.006	1.16	2
ST034Q06TA[R]	**C.** I feel lonely at school.	1	4	1.97	0.85	0.74	0.010	1.26	2
ST123Q02NA	**A.** My parents support my educational efforts & achievements	1	4	3.25	0.87	-1.23	0.032	1.83	2
ST123Q03NA	**B.** My parents support me when I am facing difficulties at school.	1	4	3.21	0.83	-1.05	0.020	1.52	2
ST123Q04NA	**C.** My parents encourage me to be confident.	1	4	3.20	0.84	-1.05	0.029	1.76	2
ST205Q01HA	**A.** Students seem to value competition.	1	4	2.55	0.85	-0.16	0.054	2.41	5
ST205Q02HA	**B.** It seems that students are competing with each other.	1	4	2.55	0.86	-0.16	0.061	2.59	5
ST205Q03HA	**C.** Students seem to share the feel that competing with each other is important	1	4	2.47	0.87	-0.09	0.057	2.49	5
ST205Q04HA	**D.** Students feel that they are being compared with others.	1	4	2.55	0.90	-0.14	0.045	2.17	5
ST062Q01TA	**A.** How often: I <skipped> a whole school day.	1	4	1.68	0.84	1.19	0.031	1.81	4
ST062Q02TA	**B.** How often: I <skipped> some classes.	1	4	1.74	0.85	1.06	0.025	1.65	4
ST062Q03TA	**C.** How often: I arrived late for school	1	4	1.70	0.84	1.13	0.041	2.07	4
ST038Q04NA	**A.** How often: Other students made fun of me.	1	4	1.45	0.81	1.75	0.015	1.39	1
ST038Q05NA	**B.** How often: I was threatened by other students.	1	4	1.34	0.74	2.16	0.049	2.28	1
ST038Q06NA	**C.** Other students took away or destroyed things that belonged to me.	1	4	1.37	0.77	2.08	0.043	2.12	1
ST038Q07NA	**D.** I got hit or pushed around by other students.	1	4	1.33	0.75	2.21	0.051	2.33	1

Note. 1= Female, 2= male
[1]Denotes scales using [1]1= Never or almost never, [1]2= A few times a year, [1]3= A few times a month, [1]4= Once a week or more,
[2]1= Strongly disagree, [2]2= Disagree [2]3= Agree, [2]4= Strongly agree;

[3]1 = Never or hardly ever, [3]2 = Some lessons, [3]3 = Most lessons, [3]4 = Every lesson;
[4]1=Never, [4]2=One/two times, [4]3=Three/four times, [4]4=Five/more times; [5]1= Not at all true, [5]2= Slightly true, [5]3= Very true, [5]4= Extremely true.
[R] The entire presented scales were reverse-coded to have appositive meaning. After reverse coding all scales have the codes from positive to negative and correspondingly the same min and max scores.

Table 3

Descriptive Statistics for Student Variables (continued]

Variable ID	description	min	Max	Mean	SD	Skew	ICC	D Eff	Codes
ST206Q01HA	**A**. How true: Students seem to value cooperation.	1	4	2.83	0.80	-0.52	0.058	2.51	5
ST206Q03HA	**B**. Students seem to share the feeling that cooperating with each other is important.	1	4	2.89	0.77	-0.52	0.057	2.49	5
ST206Q04HA	**C**. Students feel that they are encouraged to coop with others.	1	4	2.87	0.79	-0.51	0.060	2.56	5
ST207Q01HA	**A**. It irritates me when nobody defends bullied students.	1	4	2.75	0.97	-0.54	0.060	2.56	2
ST207Q02HA	**B.** It’s a good thing to help students when can't defend them	1	4	2.90	0.87	-0.75	0.053	2.38	2
ST207Q03HA	**C.** It is a wrong thing to join in bullying.	1	4	2.93	0.93	-0.72	0.070	2.82	2
ST207Q04HA	**D.** I feel bad seeing other students bullied.	1	4	2.84	0.90	-0.62	0.047	2.22	2
ST207Q05HA	**E**. I love it when somebody stands up for other students when being bullied.	1	4	2.93	0.90	-0.78	0.053	2.38	2
ESCS	Index of economic, social and cultural status	-4.6	3.99	-0.23	0.84	-0.20	0.234	7.10	
PVMMATH	Mean of PV1MATH-PV5MATH	188	741	456	84.9	0.16	0.378	10.9	
PVMSCIE	Mean of PV1SCIE-PV5SCIE	197	712	431	82.1	0.48	0.413	11.8	
PVMREAD	Mean of PV1READ-PV5READ	183	716	422	85.7	0.37	0.399	11.4	

Note. 1= Female, 2= male
[1]Denotes scales using [1]1= Never or almost never, [1]2= A few times a year, [1]3= A few times a month, [1]4= Once a week or more,
[2]1= Strongly disagree, [2]2= Disagree [2]3= Agree, [2]4= Strongly agree;
[3]1 = Never or hardly ever, [3]2 = Some lessons, [3]3 = Most lessons, [3]4 = Every lesson;
[4]1=Never, [4]2=One/two times, [4]3=Three/four times, [4]4=Five/more times;
[5]1= Not at all true, [5]2= Slightly true, [5]3= Very true, [5]4= Extremely true;
[R] The entire presented scales were reverse-coded to have appositive meaning. After reverse coding all scales have the codes from positive to negative and correspondingly the same min and max scores.

4.5 RQ2: The variation of perception of school climate-related factors within and between schools in Kazakhstan

Concerning, the variation between schools, ICC was low for all class climate factors and ranged from 1% to 7%. In this respect, the lowest ICC was for Lack of sense of belonging at school (ST034) and the highest for Anti-bullying attitude (ST207). The Design effect was also low since it correlates with ICC and was calculated directly by the usage of ICC scores. Because the intra-class correlation coefficients and design effects were quite low (e.g., ICCs were less than .10) as depicted in Table 3), multilevel modelling of the data for the measurement and structural models was not necessary. Therefore, single-level modeling was undertaken on the data. It means there is no substantive variation in school climate factors between schools. This, in itself, is an important finding suggesting that there are not systemic climate-related effects dependent on schools themselves and that student perception of school climate largely varies within each school ecology. This is an important result in itself and will be discussed in the discussion chapter of this manuscript.

Nevertheless, the dependent variable, ESCS revealed a high ICC of 0.234, which suggests students do not have equal opportunities and differ a lot in terms of social and economic background. Furthermore, all the outcomes of interest also revealed a large proportion of variance due to between-school effects. The resultant ICCs were as follows: PVMMATH= 0.378, PVMSCIE = 0.413, PVMREAD = 0.399. Overall, this finding suggests that the school that student academic outcomes are, in large part, determined by the schools that the students attend, while the experience of school climate is not determined by the school itself but individual student experiences within each school ecology.

4.6 RQ3: Measurement model that best represents the experience of class climate for Kazakhstani students

This subsection provides a description of the measurement model that best depicts student perception of school climate in Kazakhstan. Table 5 presents inter factor correlation matrix for the measurement model. The model includes ten factors, so only one factor (Student truancy and lateness (ST062) was removed from table 4 due to low-loading items and associated poor model fit with related inter-factor variance r^2 > AVE. According to the requirements of AVE-SE rule for discriminant validity, all the other factors met the criteria of AVE > 50 rule. However, Disruptive Student Academic Behavior (ST097) was just below 0.50 at 0.48. However, this factor, met the requirements where the related inter-factor variance r^2 < AVE. Therefore, the factor was retained. Concerning the other factors, the AVEs were quite high ranging from 0.61-0.80 which suggests that all factors delineate well from each other.

Overall, the correlation between factors were significant except for two relations: (1) The relation between Disruptive Student Academic Behavior ("ST097") and student competition ("ST205"); and, (2) student competition ("ST205") and lack of sense of belonging at school ("ST034") were both found to be statistically insignificant. This suggested that these two sets of factors were not related at all.

Figure 3 demonstrates the standardized item-factor loadings. All the loadings are statistically significant within in the accepted 0.4-1.0 range. Furthermore, all the *p* values for item correlations are below $p < 0.001$.

Table 4 below shows 11 school climate factors together with (Student truancy and lateness (ST062), whereas the latter table 5 depicts ten factors after removing the unfit data.

Table 4

Inter-Factor Correlation Matrix for the Measurement Model of Student-Perceived School Climate Factors (eleven factors)

	Experience of bullying r(r^2)	Anti-bullying attitude r(r^2)	Disruptive Std Academic Behavior r(r^2)	Teacher enthusiasm r(r^2)	Teacher's support and teaching practices r(r^2)	Teacher behavior and student learning r(r^2)	Student co-operation r(r^2)	Student competition r(r^2)	Lack of Sense of belonging at school r(r^2)	Parental involvement in school activities r(r^2)	Student Truancy and Lateness r(r^2)
AVE	**.78†**	**.73†**	**.48†**	**.70†**	**.57†**	**.74†**	**.77†**	**.70†**	**.61†**	**.80†**	**.52†**
Experience of bullying	**1**										
Anti-bullying attitude	**-0.077***** (0.006)	**1**									
Disruptive Std Academic Behavior	**0.176***** (0.031)	**-0.051***** (0.003)	**1**								
Teacher enthusiasm	**-0.086***** (0.007)	**0.116***** (0.013)	**-0.297***** (0.088)	**1**							
Teacher's support and teaching practices	**-0.108***** (0.012)	**0.077***** (0.006)	**-0.242***** (0.059)	**0.338***** (0.114)	**1**						
Teacher behavior and student learning	**-0.089***** (0.008)	**0.075***** (0.006)	**-0.215***** (0.046)	**0.545***** (0.297)	**0.230***** (0.053)	**1**					
Student co-operation	**-0.107***** (0.011)	**0.371***** (0.138)	**-0.211***** (0.045)	**0.280***** (0.078)	**0.246***** (0.061)	**0.206***** (0.042)	**1**				
Student competition	**0.070***** (0.005)	**0.186***** (0.035)	**0.015**^{ns} (0.000)	**0.090***** (0.008)	**0.056***** (0.003)	**0.056***** (0.003)	**0.270***** (0.073)	**1**			
Lack of Sense of belonging at school	**0.194***** (0.038)	**-0.126***** (0.016)	**0.209***** (0.044)	**-0.155***** (0.024)	**-0.157***** (0.025)	**-0.144***** (0.021)	**-0.220***** (0.048)	**-0.009**^{ns}	**1**		
Parental involvement in school activities	**-0.181***** (0.033)	**0.287***** (0.082)	**-0.122***** (0.015)	**0.235***** (0.055)	**0.158***** (0.025)	**0.199***** (0.040)	**0.317***** (0.100)	**0.200***** (0.04)	**-0.220***** (0.048)	**1**	
Student Truancy and Lateness	**0.756***** (0.572)	**-0.110**** (0.012)	**0.650***** (0.423)	**-0.264***** (0.070)	**-0.214***** (0.046)	**-0.219***** (0.048)	**-0.291***** (0.085)	**0.015**^{ns} (0.000)	**0.265***** (0.070)	**-0.240***** (0.058)	**1**

Note. ****p* < .001, ***p* < .01, **p* < .05, ^{ns}p > .05; † Factor meets requirements for AVE-SE rule for discriminant validity: AVE>50; Related inter-factor variance (r^2) < AVE

Table 5

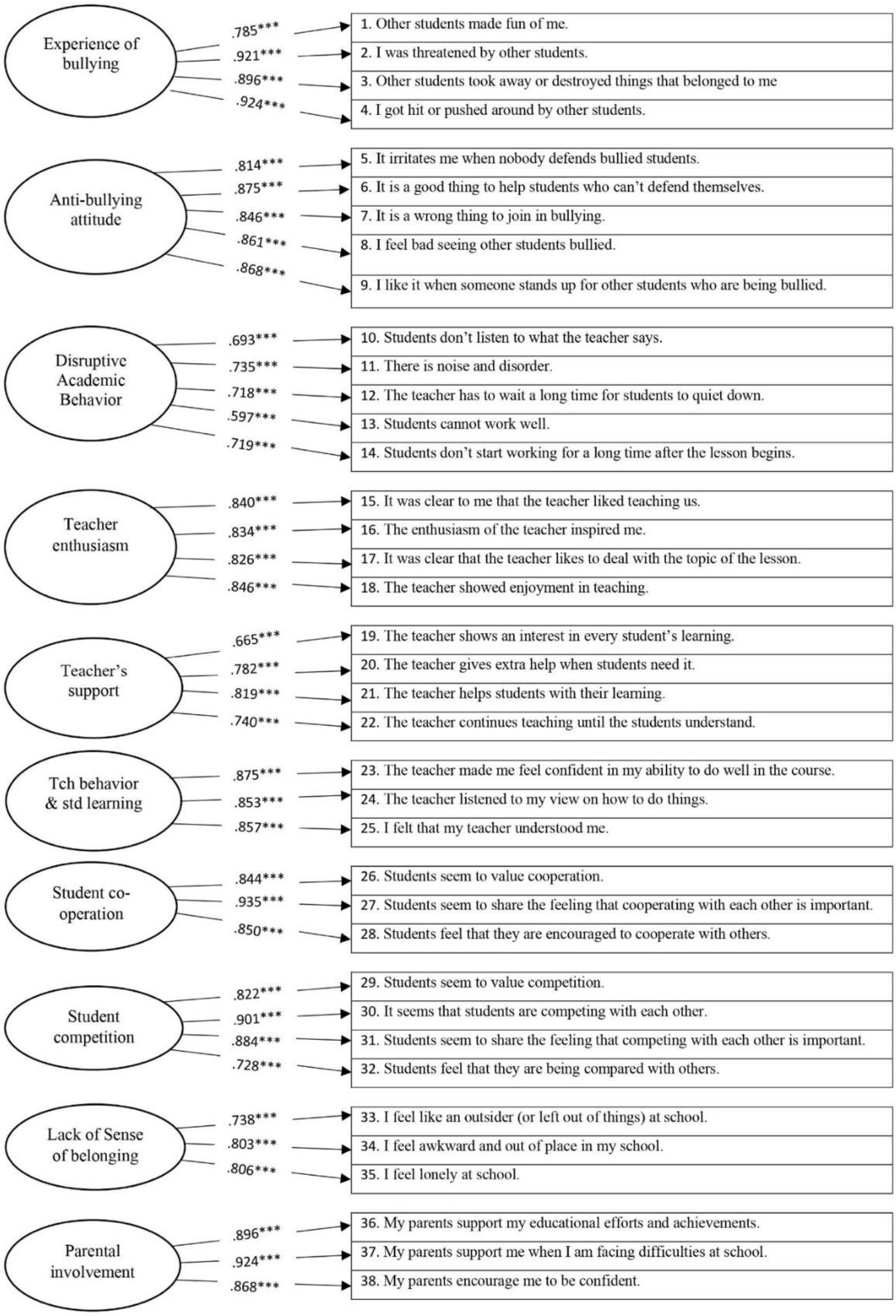
Inter-Factor Correlation Matrix for the Measurement Model of Student-Perceived School Climate Factors (ten factors)

	Experience of bullying r(r²)	Anti-bullying attitude r(r²)	Disruptive Std Academic Behavior r(r²)	Teacher enthusiasm r(r²)	Teacher's support and teaching practices r(r²)	Teacher behavior and student learning r(r²)	Student co-operation r(r²)	Student competition r(r²)	Lack of Sense of belonging at school r(r²)	Parental involvement in school activities r(r²)
AVE	**.78†**	**.73†**	**.48†**	**.70†**	**.57†**	**.74†**	**.77†**	**.70†**	**.61†**	**.80†**
Experience of bullying	**1**									
Anti-bullying attitude	**-0.076***** (0.006)	**1**								
Disruptive Std Academic Behavior	**0.176***** (0.031)	**-0.051***** (0.003)	**1**							
Teacher enthusiasm	**-0.086***** (0.007)	**0.116***** (0.013)	**-0.296***** (0.088)	**1**						
Teacher's support and teaching practices	**-0.109***** (0.012)	**0.077***** (0.006)	**-0.242***** (0.059)	**0.337***** (0.114)	**1**					
Teacher behavior and student learning	**-0.089***** (0.008)	**0.075***** (0.006)	**-0.215***** (0.046)	**0.545***** (0.297)	**0.230***** (0.053)	**1**				
Student co-operation	**-0.107***** (0.011)	**0.371***** (0.138)	**-0.210***** (0.044)	**0.280***** (0.078)	**0.246***** (0.061)	**0.206***** (0.042)	**1**			
Student competition	**0.070***** (0.005)	**0.186***** (0.035)	**0.015ns** (0.000)	**0.090***** (0.008)	**0.056***** (0.003)	**0.056***** (0.003)	**0.270***** (0.073)	**1**		
Lack of Sense of belonging at school	**0.194***** (0.038)	**-0.126***** (0.016)	**0.208***** (0.043)	**-0.155***** (0.024)	**-0.157***** (0.025)	**-0.144***** (0.021)	**-0.220***** (0.048)	**-0.009ns**	**1**	
Parental involvement in school activities	**-0.182***** (0.033)	**0.287***** (0.082)	**-0.122***** (0.015)	**0.235***** (0.055)	**0.158***** (0.025)	**0.199***** (0.040)	**0.316***** (0.100)	**0.200***** (0.04)	**-0.220***** (0.048)	**1**

Note. ****p* < .001, ***p* < .01, **p* < .05, ns p > .05; † Factor meets requirements for AVE-SE rule for discriminant validity: AVE>50; Related inter-factor variance (r^2) < AVE

Figure 3

Item-Factor Loadings for the Measurement Model of Student-Perceived School Climate Factors



Note. ****p* < .001, ***p* < .01, **p* < .05, ^{ns}p > .05

4.7 RQ4: Structural model that best represents the effect of class climate factors on student academic performance

In the structural model, all ten class climate factors were modelled as independent variables and the three forms of academic performance were modelled as dependent variables. Analysis suggested that eight class climate factors, namely, Anti-bullying attitude, Experience of bullying, Student competition, Teacher enthusiasm, Disruptive Student Academic Behavior, Lack of sense of belonging at school, Parental involvement in school activities, and Teacher's support and teaching practices were significant for all three subjects of interests. The other two factors, namely, Teacher behavior and student learning and Student co-operation were found to be insignificant for all three subjects except for Student co-operation on PVMMATH (see Figure 4). According to the results, students' Anti-bullying attitude had the most positive impact on student achievement in all three subjects with the following effects, ($b = 0.221$, $p < .001$), (b=0.313, $p < .001$), (b=0.324, $p < .001$) for Maths, Science, and Reading, respectively. The second highest overall effect appeared to be student Experience of bullying but this category, expectedly, had a negative impact at (b=-0.184, $p < .001$), (b=-0.212, $p < .001$), and (b=-0.253, $p > .001$), respectively.

The third most important class climate factor appeared to be Student competition with the highest impact on PVMMATH at ($b = .159$, $p < .001$), and with a smaller impact on PVMSCIE ($b = .092$, $p < .001$) and PVMREAD ($b = 0.072$, $p < .001$). The fourth category was Teacher enthusiasm with a negative impact. However, the effects were quite small in terms of practical significance at ($b = -0.066$, $p < .001$),(b= -0.079, $p < .001$) and (b=-0.090, $p > .05$.) The next category, Disruptive Student Academic Behavior (ST097) also had the negative impact on student academic performance, but it was expected in comparison to the former factor. The effects were also quite small at ($b = -0.050$, $p < .001$), (b=-0.042, $p < .001$), and (b= -0.038, $p < .001$). A Lack of sense of belonging at school was another anticipated value with a negative effect. The results were at (b = -0.030, $p < .001$), (b=-0.029, $p < .001$) and (b=-0.043, $p < .001$) respectively.

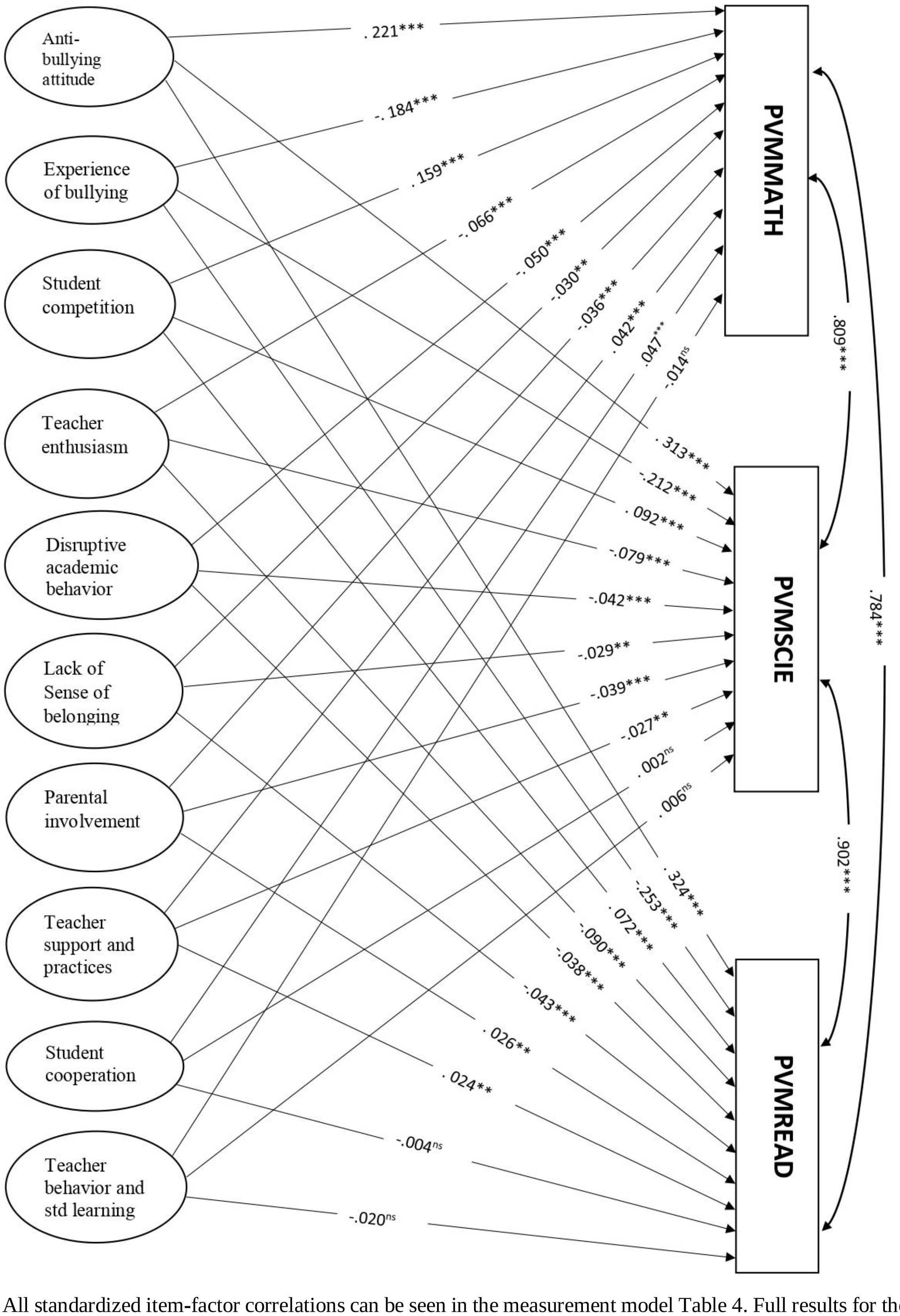
Parental involvement in school activities had a small negative impact on PVMMATH (b = -0.36, $p < .001$) and PVMSCIE (b = -0.039, $p < .001$), whereas a negligible positive impact on PVMREAD (b = 0.026). Teacher's support and teaching practices was the least important factor from the statistically significant categories of factors at (b=0.042, $p < .001$) for PVMMATH, (b = -0.027, $p < .001$) for PVMSCIE, and (b = 0.024, $p < .001$) for PVMREAD. Teacher behavior and student learning was insignificant for all outcomes at (b=-0.014, $p > .05$), (b=0.006, $p > .05$) and (b=-0.020, $p > .05$). Student co-operation was significant only for PVMMATH (b = 0.047, $p < .001$), wheras the effect on PVMSCIE (b = 0.002, $p > .05$) and PVMREAD, (b = -0.004, $p > .05$) were found to be insignificant (see Figure 4).

As for the relationship among the outcomes, all of them were found to be highly and statistically, and significantly correlated. In this case, the correlation between PVMMATH and PVMSCIE, was (b=0.809, $p < .001$), while for PVMMATH and PVMREAD, it was (b=0.784, $p < .001$), and the highest correlation was between PVMSCIE and PVMREAD at (b=0.902, $p < .001$).

Lastly, the R^2 values for all three dependent variables were estimated. This reflects the total proportion of variance explained in the dependent variables for PVMMATH (R^2 = .141), PVMSCIE (R^2 = .167), PVMREAD (R^2 = .196). Further, the R^2 values were used to identify f^2 effect size values (Cohen, 1992). According to the results, it was found that the class climate factors all have a medium effect on students' mathematics, science, and reading scores. For PVMMATH, f^2 = 0.16, for PVMSCIE, f^2 = 0.20, and for PVMREAD, f^2 = 0.24. To conclude it was identified that school climate factors have the most substantive effect on Kazakhstani student reading literacy**.**

Figure 4

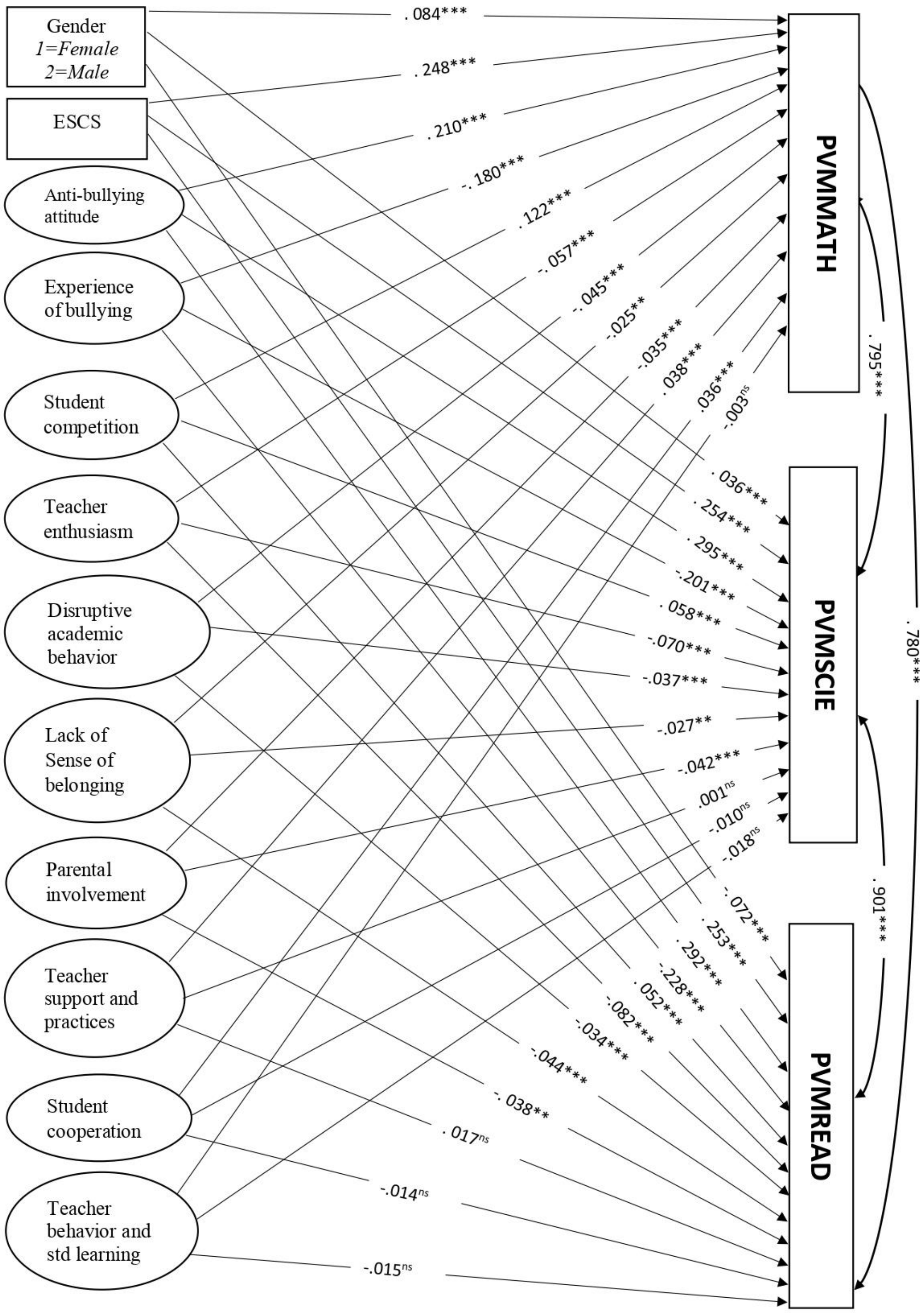
Model for Class Climate Factors on Student Academic Performance in Kazakhstan



Note. All standardized item-factor correlations can be seen in the measurement model Table 4. Full results for the structural model are available from the author; ***Note***. ***p < .001, **p < .01, *p < .05, [ns]p > .05.

Figure 5

Model for Class Climate Factors and demographic variables on Student Academic Performance in Kazakhstan



Note. Full results for the structural model with Gender and ESCS are available from the author; ***Note***. *** $p < .001$, ** $p < .01$, * $p < .05$, $^{ns}p > .05$. Values equal to .001 were mentioned as *** as well.

After including, Gender and ESCS in the structural model the following results were taken. An anti-bullying attitude remained the most important driver of student achievement in all three subjects with the following effects, (b = 0. 210, p < .001), (b=0. 295, p < .001), (b=0.292, p < .001) for Maths, Science, and Reading, respectively. Its effect was bigger than the effect of ESCS, which only had more impact on Math (b=0.248, p < .001), while the influence on Science (b = 0. 254, p < .001), and Reading (b=0. 253, p < .001) were smaller than the former. The second highest overall effect on Maths, Science, and Reading remained unchanged with student Experience of bullying at (b= -0.180, p < .001), (b=-0.201, p < .001), and (b=-0.228, p <.001) respectively. It can be noted that after incorporating demographic variables all school climate factors, both with positive and negative effect had slightly less impact on math. As for the Science impact, all factors had a weaker positive or negative impact except for Parental involvement (b = -0.042, p < .001) which had a slightly more negative impact. Furthermore, one more factor (Teacher's support and teaching practices) was found insignificant (b=0.001, p > .05) which was previously moderately significant. Concerning the Science results, all categories had weaker positive or negative effects except for A Lack of sense of belonging at school (b=-0.044, p < .001) with a slightly stronger negative effect. The only factor which changed its direction was Parental involvement (b=-0.038, p < .01) which became moderately negative despite being moderately positive prior. As in the Reading effect Teacher's support and teaching practices also became insignificant for Science. Regarding the Gender impact, it can be concluded that boys were found better in Math and Science, whereas girls were better in Reading (see Figure 5).

Finally, the R^2 values for all three dependent variables were reestimated together with demographic variables. This time there were the following R^2 values PVMMATH (R^2 = .184), PVMSCIE (R^2 = .204), PVMREAD (R^2 = .223). After calculating the f^2 effect size values (Cohen, 1992) with Gender and ESCS it can be concluded that the effects on students' mathematics, science, and reading scores increased. For PVMMATH, f^2 = 0.23, for PVMSCIE, f^2 = 0.26, and for PVMREAD, f^2 = 0.29. To conclude it was identified that school climate factors have a stronger impact while including gender and ESCS and the most substantive effect on Kazakhstani students remained reading literacy.

4.8 Model fit

All model fit indices were interpreted in accordance with the methodology. The χ^2/df ratio was 5,440.940/704, i.e., 7.73, which was statistically significant (p < .001). However, this result is quite common with larger models and large numbers of students. Other less biased fit indices were also examined. Comparative fit index and Tucker-Lewis Index both produced the same number and exceeded the accepted values of .90 with estimates of CFI = .98, TLI = .98. Finally RMSEA = .027, with (L=0.026 and U=0.028). After including Gender and ESCS in the structural model, the following changes occurred. The χ^2/df ratio grew and it was 6841.238/780, i.e., 8.77, with TLI = .97, slightly higher SRMR = .034 and RMSEA = .029, with (L=0.028 and U=0.030), whereas p value and CFI stayed unchanged. Three decimal places were selected for this category due to the numbers being close to each other (see Table 6).

Table 6

Model Fit Indices for Measurement and Structural Model

	Model Fit Indices								
Model	*N*	X^2	*df*	X^2/df	*P*	CFI	TLI	SRMR	RMSEA
CFA & SEM	11,317	5,440.940	704	7.73	< .001	.98	.98	.023	.027(L=.026, U =.028)
SEM with Gender and ESCS	11317	6841.238	780	8.77	< .001	.98	.97	.034	.029(L=.028, U =.030)

Note. *N* = number of observations; χ^2 = Chi-squared; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual; RMSEA = root mean square error of approximation; U = upper 90% confidence interval; L=lower 90% confidence interval.

Therefore, an inspection of model fit suggested that the model was a sound representation of the data and, therefore, the model may provide a substantive contribution to the theory concerning the role of class climate on student academic performance in Kazakhstan.

4.9 Summary of the findings

This chapter introduced the major findings of the study on school climate. The results were outlined regarding the four research questions. The data was described in accordance with survey questions on secondary data. Results suggested that a reasonable level of school climate exists for students in Kazakhstan, which is not much different from the OECD average. The biggest differences occurred with Disruptive Student Academic Behavior and Student truancy. In this regard, Kazakhstani students seemed to have a much more pleasant undistracted classroom environment than the other OECD countries, whereas Student truancy was the major issue for Kazakhstani students in comparison with others (OECD, 2019a). However, results also suggested that school climate does not vary between schools in any systemic way. The OECD scales appear to function in a valid way to measure student perception of school climate in Kazakhstan. In terms of the main research question, RQ4, it was found that only a few factors were found the drivers of student performance, which somehow contradicts the literature review which identified that all of the factors were of importance, the findings themselves generally support the OECD (2019) average among all countries. Therefore, it can be argued that the findings of this study were surprising and expected at the same time. The next chapter provides a discussion of the results of the research with respect to the reported literature review.

5. Discussion

5.1 Introduction

This chapter elaborates on the results of the current study tasked with determining how student perception of school climate is related to Kazakhstani students' academic performance in PISA. This chapter follows four research questions and covers five sections. The first section discusses the relative level of school climate in Kazakhstan, the second section discusses the degree to which student-level school climate varies within- and between-schools in Kazakhstan, the third section covers the measurement model that best represents the experience of class climate for Kazakhstani students. Finally, the fourth section discusses the major findings in this study, the school-climate related factors that drive student academic outcomes in the country. Finally, the chapter is summarized by way of the conclusion of the main ideas.

5.2 RQ1: Prominent student school climate factors in Kazakhstan

According to students' responses on school climate, the following ideas can be presented. It was found out that Kazakhstani students generally tend to obey the teacher and that there is generally not much noise and disarrangement in the classroom. Additionally, teachers did not seem to wait a lot to create a calm environment and students do not seem to face issues with working well and engaging once the lesson begins. Since the lessons are not usually disrupted by the students, this category of school climate appears to be quite sound in comparison with international averages, where classroom disruption likely to be administered three times more (OECD, 2019a).

The level of reported teacher support also tended to be quite positive, which means teachers tend to show sympathy for every student's learning and give additional assistance to students when needed. Students state that teachers support their learning and proceed with teaching until students' full comprehension. Overall, teachers were found to be supportive, and students were satisfied with their teachers' assistance. The majority of students claim that their parents are also supportive of their educational attempts and accomplishments. Parents were found to be supportive when students face difficulties at school. Furthermore, students claim that their parents support them to be more confident.

Furthermore, most students found their teacher to appreciate them. Students felt they were inspired by the enthusiasm of their teachers. Students also state teachers tend to have both love for the topic of the lesson and teaching.

However, the students did not value competition very highly. When compared with cooperation, which is usually linked and correlated with competition, cooperation was valued more. Attendance seemed to cause twice more issues for Kazakhstani students than to other OECD countries, but arriving late to school did not show much difference (OECD, 2019a). Furthermore, student safety seems to cause some problems. The issues such as mockery, threatening, destroying or taking away belongings, and hitting or pushing were reported several times a year. Though the results did not reveal an abundance of student bullying, it can be concluded that being bullied several times a year is common for the regular student. As for the anti-bullying attitude, students stated they generally feel irritated to see the suffering of bullied students and like it when somebody stands up for them. However, the results were slightly less than in OECD countries. Most of the students acknowledged that bullying was maladaptive. It appeared fortunate that students seem to disapprove of such deviant behavior and have a strong anti-bullying attitude.

5.3 RQ2: The variation of student-level school climate-related factors

Student-level school climate factors have a good variation within schools and did not seem to vary much between schools as evidenced by the low ICC estimates. It can be concluded that the majority of schools are not likely to differ systematically in terms of school belongingness. This low variation can be explained with the OECD (2019a) report which states a sense of belonging varies relating to the location: rural or city and the school type: public or private.

Therefore, it can be concluded the number of public schools from the city prevailed in the current study. Concerning students' responses, there were not many differences from the OECD average. For example, if in Kazakhstan (counting only participants of the study) 79% of students noted they feel lonely at school, whereas for OECD it was 84%. Regarding students feeling like outsiders the results almost equaled with 79% and 80% to the slight advantage of OECD countries. Another lowest school variation was for perceived support from parents. Despite most students reported being provided strong support from parents, composite scores did not differ across schools.

Comparatively, students' Anti-bullying Attitude and experience of disruptive student behavior have slightly wider variation at the school level. In this regard, there might be differences in school policy and class climate. Student cooperation and competition also exhibited slightly more between-school variance indicating the level of students' engagement in some schools is diverse. However, students' responses to different questions within Teacher Enthusiasm and Student Lateness and Truancy scale were unstable and had a misbalance of answers despite being within a set of factors. Therefore, it was difficult to draw conclusions based on these findings. It is supported by the fact Student Lateness and Truancy were found unfit and were removed from the final structural model, whereas Teacher Enthusiasm had an only unsubstantive effect.

According to students' scores on the test, there was a lot of variation for all outcomes of interest (mathematics, science, and reading). The same high variation was present in student socio-economic status. These statistics reveal that SES has a huge influence on students' academic attainment in PISA. This finding supports previous studies that students with higher SES are predisposed to be more successful at school (Gilleece, Cosgrove, & Sofroniou, 2010; Chen, 2016).

5.4 RQ3: Measurement model that best represents the experience of class climate for Kazakhstani students.

A ten-factor measurement model for student-perceived school climate factors was specified to represent the observed data. The measurement model was good-fitting and supported construct validity.

All inter-factor correlations were found to be statistically significant except for two relationships. These were the correlations between (1) Disruptive Student Academic Behavior (ST097) and Student competition (ST205), and (2) Student competition (ST205) and Lack of sense of belonging at school (ST034).

The size and direction of the inter-factor correlations suggest that students gave authentic answers in general support of the theory. Correlations were generally quite small suggesting that each factor was distinct. Only one inter-factor correlation could be considered large, that is, the relationship between ‘Teacher Enthusiasm’ and ‘Teacher Behaviour and Student Learning’ ($r = .545$, $p < .001$). This is in general support of other findings in the field. According to the findings, expressive and explicit teachers tended to be successful with their students learning facts, while energized teachers were likely to increase their students’ comprehension (Solomon, Rosenberg, & Bezdek, 1964). Furthermore, enthusiastic teachers were better at enhancing students’ academic achievements (MacNeil et al., 2009; Kalyar et al., 2018).

There was also a moderately sized correlation observed between ‘Teacher Enthusiasm’ and ‘Teacher’s Support and Teaching Practices’ ($r = .337$, $p < .001$). This result is also supported by previous research in the field (Frenzel et al., 2009; Kalyar et al., 2018), which supports that motivated teachers tended to be more proactive by resulting in student enjoyment.

Finally, the only other moderate correlations was found between students' perception of 'Parental Involvement in School Activities' and 'Student Cooperation' (r = .316, p < .001). This finding is interesting because it suggests that the role of parents is essential to develop students' collaborative skills. (Interpretation of size of correlations made in accordance with Cohen, 1992; small r > .10; medium r > .30 and large r > .50).

The main result for RQ3 is that the specified measurement model provides a well-fitting representation of the observed data and the results of the current study can be generalised at the broader target 15-year-old population in Kazakhstan.

5.5 RQ4: The major drivers of student academic outcomes

The Structural model identifies the major drivers of student academic outcomes. Specifically, the role of school climate-related factors will be discussed in the general order with which they explain variance in the academic outcomes of interest. This research finds students' anti-bullying attitude as the school-climate related predictor for student performance. It supports findings of the OECD (2019a). This finding supports the notion that education alone is not sufficient for improving students' behavior. In this regard, the importance of positive thinking and pro-social behavior appears to buttress student learning. Therefore, it can be concluded that positive thinking and the disapproval of violence may not only support students socially but also enable them to achieve more positive academic results. To conclude, it means all education stakeholders in Kazakhstani high schools should actively participate in building an anti-bullying attitude among students.

Overall, the second most important factor for student performance for all subjects pertained to student Experience of Bullying. The results also support findings presented in various other investigations (OECD, 2019a; Eccles, 2008; Kutsyuruba et al., 2015; Nansel et al., 2001). This theory can be applied to Maslow's Hierarchy of Needs, as safety holds a fundamental position in the hierarchy. The findings also support Eccles (2008) who rated safety as an important factor of student development and studiousness. Therefore, we can observe that both the literature review and the findings of this research indicate that the experience of bullying is an important driver of (negative) student academic outcomes.

Student competition can be considered as the third most important factor for PISA academic performance. However, the effect only applies to Kazakhstani student math performance, since its influence on reading and science was low and cannot be categorized as a driver factor ($b < .10$).

The same rule applies to the remaining factors, the influence of Teacher enthusiasm, Disruptive academic behavior, Lack of Sense of belonging, Parental involvement, and Teacher support and practices were discovered as statistically significant. However, the results can not be generalized due to the low beta value. As for the other two factors, 'Teacher Behavior and Student Learning' and 'Student Cooperation', both were insignificant for all student academic subjects except for the small influence of Student Cooperation on mathematics. When including Gender and ESCS the number of insignificant values increased to three. For example, 'Teacher Support and Practices', 'Teacher Behavior and Student Learning' and 'Student Cooperation' were all found to be insignificant for science and reading scores. In this regard, 'Teacher Behavior and Student Learning' was the only factor with insignificant effect for all three subjects, whereas 'Teacher Support and Practices' and 'Student Cooperation' were positively significant for math performance. These results somehow contradict previous studies on the importance of 'Teacher Behavior and Student Learning' on student academic performance. The following conclusions can be drawn. Kyriakides et al. (2009) stated that as teacher behavior (according to five types of teacher behavior) advances so does student behavior increases. In this regard, the initial types of teacher behavior incorporate basic, direct teaching with some elements of active teaching, while the advanced category includes more active teaching, new teaching approaches, and differentiation (Kyriakides et al.). Therefore, according to the findings, it can be concluded Kazakhstani teachers tended to put more effort into direct and basic teaching, whereas new teaching approaches and including elements of differentiation seems to be lacking.

5.6 Summary

The discussion section elaborated on all four research questions and provided corresponding explanations behind the findings.

The findings of this study both support and run against findings in previous studies. For example, two out of ten school climate factors were the important predictors of student achievement in all three subjects of interest. However, most of the factors almost did not affect student PISA performance despite being statistically significant. Out of all categories two factors, specifically ‘Teacher Behavior and Learning’ and ‘Student Cooperation’ were found to be largely insignificant for student academic outcomes. After incorporating gender and ESCS Teacher Support and Practices also became insignificant. To conclude, not all school climate factors important for student achievement, and there two leading factors and two to three least important categories.

6. Conclusion and Implications

6.1 Summary of research findings

The purpose of this quantitative study was to examine the influence of school climate-related factors on students’ academic outcomes in Kazakhstan. It aimed to measure the effect of school climate factors on student academic achievements namely, Math, Reading, Science results. The summary of the research findings is presented by revisiting the research questions.

Student level variables were presented by descriptive statistics which states that the majority of the students found their teachers to be supportive and enthusiastic. Students did not seem to face issues with classroom noise and parental involvement. In this regard, the noise level in the Kazakhstani classroom was three times less than in OECD (2019a) average at 26 %. Students rated cooperation above the competition and reported some lack of belongingness to the school and experience of bullying. However, it was important that bullying was not supported by the majority of the students anyway Kazakhstani students were found less compassionate than other OECD countries in terms of anti-bullying attitude.

Importantly, the findings of the study suggest that a student experience of school climate is largely individual and not systemic to particular schools or school systems. Therefore, it seems schools in Kazakhstan have quite similar composite school climate scores despite variation within schools. However, despite low ICC scores existing variation had some range from 1% to 7 % among factors. Out of all school climate factors, the lowest variation was noticed by Lack of Sense of belonging which was only 1%, whereas an Anti-bullying Attitude depicted the highest variation with 7%. Alongside, school climate categories some demographic variables, Gender, and ESCS, as well as the outcomes of interest PVMMATH, PVMSCIE, and PVMREAD were also included in the descriptive statistics. In this matter, schools approximately had similar small gender variation, whereas socio-economic status and scores for outcomes of interest had strong variation. The reasons behind this finding will be discussed and summarized while elaborating on the main research question.

A single-level ten-factor measurement model for student experience of school climate represents a useful way to conceptualize students' experience of the school social and academic environment in Kazakhstan. Results were generally in support of previous theory but also extended by way of identifying a relation between parental involvement and student cooperation. Previously, the role of parents and student motivation, and student academic performance were researched. However, the relation between parental involvement and student cooperation can be viewed as a new concept. Therefore, this finding seems to be crucial.

Although the majority of the school climate factors were significant, this research has established that out of all school climate factors the most important drivers of student performance in PISA were Experience of bullying and Anti-bullying attitude. Concerning the driver factors, it supports previous theories on having an anti-bullying attitude OECD (2019a) and a safe environment for learning (OECD; Eccles, 2008; Kutsyuruba et al., 2015; Nansel et al., 2001) and paying attention to cases of bullying at school. However, 'Teacher Support and Practices', 'Teacher Behavior and Student Learning' and 'Student Cooperation' were not found important, since the research has established these factors as insignificant. Furthermore, according to demographic variables, the following conclusions can be drawn. Students with higher socio-economic status have more chances to have better scores in all three subjects: Math, Science, and Reading. Regarding Gender, boys demonstrated better results in Math and Science, while girls scored higher in Reading. To conclude, boys who are not bullied, having a strong Anti-bullying attitude and SES were likely to succeed in Math and Science, whereas equipped with all the aforementioned factors girls tend to be successful with Reading.

6.2 Limitations and Recommendations

As for the recommendations in the upcoming studies, researchers may look at the influence of school climate factors on other educationally-related outcomes such as re-enrolment, attendance, and participation in extra-curricular activities--other important outcomes in school. Furthermore, the researchers may analyze the disruptive student behavior deeper since it was discovered as the most crucial school climate factor. Its relation with other scales may be studied, too. The relation between Teacher Enthusiasm and Teacher Behaviour and Student Learning could also be researched since the current study found a strong relationship between these categories. Therefore, there are many ways how this type of research can be elaborated and extended in the future.

6.3 Recommendation for future research

It is recommended to conduct this research for future PISA administrations in order to confirm these initial results. Future research could also employ comparative work in the Central Asian and post-Soviet block region. Since this research covered school climate from students' perspectives, future work in this field might also incorporate teachers, parents, and administrators. In fact, future investigations may look at how stakeholders' perceptions also contribute to student academic performance. School climate is a rich topic and there are plenty of opportunities out there for future researchers.

6.3.1 Recommendation for administration

School climate is an important indicator for not only improving student achievements but also for improving the quality of the school as well. The literature reviewed suggests that administrators are the main figures in establishing a school climate (Deal & Peterson, 1998; Frost et al., 2014). Therefore, their roles are crucial for the development of the school. Concerning the findings specific to this research, school administrators should pay attention to school safety by recognizing and decreasing the number of bullying incidents at school. The administrators should not only work arduously toward such targets with the safety of students in mind. Furthermore, regarding the findings, building an anti-bullying attitude is highly important. Therefore, additional professionally run classes and instruction should be organized for these types of lessons. Implementation of such programs can be monitored and the effect of such interventions evaluated with the outcome focused on increasing pro-social behavior and, ultimately, improved academic performance. In this matter, it can be concluded that building an anti-bullying attitude and creating a safe environment for the study is essential.

6.3.2 Recommendation for Ministry of Education

Specialized training courses could be developed and run by the Ministry of Education for teachers, school leaders, and administrators. These courses may be organized both at school (for in-service staff) and at teacher training centers. The content of the course shall incorporate the importance of (1) anti-bullying attitudes, and (2) the need to build a friendly atmosphere by excluding bullying incidents. Further, the significance of the above-mentioned theories may be mentioned during lessons. Furthermore, the Ministry of Education may organize or approve initiatives to run special clubs for building anti-bullying attitudes as initiated by students.

7. References

Ab Hamid, M. R., Sami, W., & Mohmad Sidek, M. H. (2017). Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. *Journal of Physics: Conference Series, 890,* 012163. https://doi.org/10.1088/1742-6596/890/1/012163

Allen, K., Kern, M. L., Vella-Brodrick, D., Hattie, J., & Waters, L. (2016). What Schools Need to Know About Fostering School Belonging: a Meta-analysis. *Educational Psychology Review*, *30*(1), 1–34. https://doi.org/10.1007/s10648-016-9389-8

Agasisti, T. (2011). Does Competition Affect Schools' Performance? Evidence from Italy through OECD-PISA Data. *European Journal of Education*, *46*(4), 549–565. https://doi.org/10.1111/j.1465-3435.2011.01500.x

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Aucejo, E. M., & Romano, T. F. (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review*, *55*, 70–87. https://doi.org/10.1016/j.econedurev.2016.08.007

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singman, H.,… Fox, J. (2020). *Lmer Performance Tips*. https://cran.r project.org/web/packages/lme4/vignettes/lmerperf.html

Benner, A. D., Boyle, A. E., & Sadler, S. (2016). Parental Involvement and Adolescents' Educational Success: The Roles of Prior Achievement and Socioeconomic Status.

Journal of Youth and Adolescence, *45*(6), 1053–1064. https://doi.org/10.1007/s10964-016-0431-4

Blank, C., & Shavit, Y. (2016). The Association Between Student Reports of Classmates' Disruptive Behavior and Student Achievement. *AERA Open*, *2*(3), 233285841665392. https://doi.org/10.1177/2332858416653921

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Browne, M., & Cudeck, R. (1989). Single Sample Cross-Validation Indices for Covariance Structures. *Multivariate Behavioral Research*, *24*(4), 445–455. https://doi.org/10.1207/s15327906mbr2404_4

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Byrne, B. M. (2001). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming (Multivariate Applications Series)* (1st ed.). Psychology Press.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational Research Review, 14*, 33–46. https://doi.org/10.1016/j.edurev.2015.01.002

Chen, Q. (2016). A Multilevel Analysis of Singaporean Students' Mathematics Performance in PISA 2012. *What Can PISA 2012 Data Tell Us?*, 17–33. https://doi.org/10.1007/978-94-6300-468-8_2

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J., McCabe, E. M., Michelli, N. M., & Pickeral, T. (2009). *School climate: Research, policy, practice, and teacher education.* Teacher College Record, 111 (1), 180-213

Creswell, J. W. (2014). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Eccles, J. S. (2008). *Can middle school reform increase high school graduation rates*. Santa Barbara, CA: Univerity of California.

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

Frase, L. E., & Sorenson, L. (1992). Teacher Motivation and Satisfaction: Impact on Participatory Management. *NASSP Bulletin*, *76*(540), 37–43. https://doi.org/10.1177/019263659207654007

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, *101*(3), 705–716. https://doi.org/10.1037/a0014695

Frost, D., Fimyar, O., Yakavets, N., & Bilyalov, D. (2014). *The role of school director in educational reform in Kazakhstan: In Education Reform and Internationalisation.* Cambridge University Press.

Georgiou, S. N., Christou, C., Stavrinides, P., & Panaoura, G. (2002). Teacher attributions of student failure and teacher behavior toward the failing student. *Psychology in the Schools*, *39*(5), 583–595. https://doi.org/10.1002/pits.10049

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. https://doi.org/10.1007/s10763-010-9199-2

Goldstein, H. (2003). Multilevel Modelling of Educational Data. *Methodology and Epistemology of Multilevel Analysis*, 25–42. https://doi.org/10.1007/978-1-4020-4675-9_2

Informational and Analytical Centre (2017). *National Report on education*. Ministry of Education and Science of the Republic of Kazakhstan.

Hampden-Thompson, G., Guzman, L., & Lippman, L. (2013). A cross-national analysis of parental involvement and student literacy. *International Journal of Comparative Sociology*, *54*(3), 246–266. https://doi.org/10.1177/0020715213501183

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1-15). Sage Publications.

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

Huebner, E. S., & Gilman, R. (2006). Students Who Like and Dislike School. *Applied Research in Quality of Life*, *1*(2), 139–150. https://doi.org/10.1007/s11482-006-9001-3

Ispambetova, B. (2018, June 4). Research ethics review as one way to protect human subject and not only. *Nazarbayev University Repository*. https://nur.nu.edu.kz/handle/123456789/3285

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin, 89*(1), 47–62. https://doi.org/10.1037/0033-2909.89.1.47

Nash, J. C., & Varadhan, R. (2011). Unifying Optimization Algorithms to Aid Software System Users: optimx for R. *Journal of Statistical Software*, *43*(9). https://doi.org/10.18637/jss.v043.i09

Nash, J. C. (2014). On Best Practice Optimization Methods inR. *Journal of Statistical Software*, *60*(2), 1-14. https://doi.org/10.18637/jss.v060.i02

Kalyar, M. N., Ahmad, B., & Kalyar, H. (2018). Does Teacher Motivation Lead to Student Motivation? The Mediating Role of Teaching Behavior. *Voprosy Obrazovaniya / Educational Studies Moscow, 3*, 91–119. https://doi.org/10.17323/1814-9545-2018-3-91-119

Klein, S. S. (1971). Student Influence on Teacher Behavior. *American Educational Research Journal*, *8*(3), 403–421. https://doi.org/10.3102/00028312008003403

Klem, A. M., & Connell, J. P. (2004). Relationships Matter: Linking Teacher Support to Student Engagement and Achievement. *Journal of School Health, 74*(7), 262–273. https://doi.org/10.1111/j.1746-1561.2004.tb08283.x

Kunter, M. (2013). Motivation as an aspect of professional competence: Research findings on teacher enthusiasm. In *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 273-289). Boston.

Kutsyuruba, B., Klinger, D. A., & Hussain, A. (2015). Relationships among school climate, school safety, and student achievement and well-being: a review of the literature. *Review of Education*, *3*(2), 103–135. https://doi.org/10.1002/rev3.3043

Kyriakides, L., Creemers, B., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, *25*(1), 12–23. https://doi.org/10.1016/j.tate.2008.06.001

Lai, M. H. C., & Kwok, O. M. (2014). Examining the Rule of Thumb of Not Using Multilevel Modeling: The "Design Effect Smaller Than Two" Rule. *The Journal of Experimental Education*, *83*(3), 423–438. https://doi.org/10.1080/00220973.2014.907229

Ma, X., & Willms, J. D. (2004). School disciplinary climate: Characteristics and effects on eighth grade achievement. *Alberta Journal of Educational Research*, *50*(2).

MacNeil, A. J., Prater, D. L., & Busch, S. (2009). The effects of school culture and climate on student achievement. *International Journal of Leadership in Education*, *12*(1), 73–84. https://doi.org/10.1080/13603120701576241

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2020). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-4.https://CRAN.R-project.org/package=e1071

MoES (Ministry of Education and Science of the Republic of Kazakhstan). 2016. Strategic Plan of Ministry of Education and Science of Kazakhstan for 2017-2021, Astana.

Muratkyzy, A. (2020, August 17). Equity and excellence in the Kazakhstani education system: A multilevel analysis on the personal and contextual factors contributing to students' reading literacy performance on PISA 2018 (Doctoral dissertation). *Nazarbayev*

University Repository. https://nur.nu.edu.kz/handle/123456789/4886

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Jama*, *285*(16), 2094-2100. http://dx.doi.org/10.1001/jama.285.16.2094.

Ning, B., Van Damme, J., Van Den Noortgate, W., Yang, X., & Gielen, S. (2015). The influence of classroom disciplinary climate of schools on reading achievement: a cross-country comparative study. *School Effectiveness and School Improvement*, *26*(4), 586–611. https://doi.org/10.1080/09243453.2015.1025796

OECD. (2014). PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science, PISA, Paris.

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

OECD. (2018). Education Policy Outlook: Kazakhstan. http://www.oecd.org/education/Education-Policy-Outlook-Country-Profile-Kazakhstan-2018.pdf.

OECD. (2019a). *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*. Paris. https://doi.org/10.1787/acd78851-en.

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

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

Peterson, K. D., & Deal, T. E. (1998). How leaders influence the culture of schools. *Educational leadership*, *56*, 28-31.

PISA 2018 Database. (n.d.). https://www.oecd.org/pisa/data/2018database/

Randolph, J. (2009). A guide to writing the dissertation literature review. *Practical Assessment, Research, and Evaluation, 14*(1), 13. DOI: https://doi.org/10.7275/b0az-8t74

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Roby, D. E. (2004). Research on school attendance and student achievement: A study of Ohio schools. *Educational Research Quarterly, 28*(1), 3-16.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

Sherman, T. M., & Cormier, W. H. (1974). An investigation of the influence of student behavior on teacher behavior. *Journal of Applied Behavior Analysis, 7*(1), 11-21.

Shindler, J., Jones, A., Williams, A. D., Taylor, C., & Cardenas, H. (2016). The School Climate-Student Achievement Connection: If We Want Achievement Gains, We Need to Begin by Improving the Climate. *Journal of School Administration Research and Development*, *1*(1), 9-16.

Solomon, D., Rosenberg, L., & Bezdek, W. E. (1964). Teacher behavior and student learning. *Journal of Educational Psychology, 55(1), 23–30.* https://doi:10.1037/h0040516

Tasbulatova, S., Belosludtseva, V., & Grooting, P. (2007). Skills development and poverty reduction in Kazakhstan. ETF working document.

Tauer, J. M., & Harackiewicz, J. M. (2004). The Effects of Cooperation and Competition on Intrinsic Motivation and Performance. *Journal of Personality and Social Psychology,* 86(6), *849–861.* https://doi:10.1037/0022-3514.86.6.849

Thompson, B. (2006). Foundations of behavioral statistics: An insight-based approach.The Guilford.

Walker, J. (2013). *Calculate probability from* x^2 *and d*. Retrieved October 11, 2012 from http://www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html

Wang, M. & Degol, J (2016), “School climate: A review of the construct, measurement, and impact on student outcomes”, *Educational Psychology Review*, 28(2), (pp. 315-352), http://dx.doi.org/10.1007/s10648-015-9319-1.

Wang, M.-T., & Sheikh-Khalil, S. (2013). *Does Parental Involvement Matter for Student Achievement and Mental Health in High School? Child Development, 85(2), 610–625.* https://doi:10.1111/cdev.12153

Weiss, C. C., & García, E. (2015). Student engagement and academic performance in Mexico: Evidence and puzzles from PISA. *Comparative education review*, *59*(2), 305-331.

Wickham, H, & Miller, E. (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1.

Yakavets, I. (2014). *Educational reform and internationalisation: The first decade of Independence*. D. Bridges (Ed.). Cambridge University Press

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

Yu, C. Y., & Muthén, B. (2002). *Evaluation of model fit indices for latent variable models with categorical and countinuous outcomes*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans.

Appendix

```
###################### R Script for Data and Analysis ###################################
############################# PISA 2018 ############################################
#### Install packages necessary for running script ####
# haven for reading SPSS .sav file
if(!require("haven")){
  install.packages("haven", dependencies=T)
  library(haven)
}
citation("haven")
# e1071 for estimating variable skewness
if(!require("e1071")){
  install.packages("e1071", dependencies=T)
  library(e1071)
}
citation("e1071")
# lme4 for running multilevel (linear mixed) models
if(!require("lme4")){
  install.packages("lme4", dependencies=T)
  library(lme4)
}
citation("lme4")
# optimx for supporting lme4 package
if(!require("optimx")){
  install.packages("optimx", dependencies=T)
  library(optimx)
}
citation("optimx")
# lavaan package
if(!require("lavaan")){
  install.packages("lavaan", dependencies=T)
  library(lavaan)
```

```
}
citation("lavaan")
#### Set working directory ####
setwd("C:/Users/Admin/Desktop/For coding/SCH")
getwd()
rm(list=ls())
# Get PISA school data
school.std.kaz.merge <- read.csv("school.std.kaz.merge.csv")
########################Student-Level Independent Variables from the PISA Dataset [Instrument 1]########
## Explore df
dim(school.std.kaz.merge) # 19507, 1313
head(school.std.kaz.merge[,1:10])
################ Create mean ESCS variable for each school ############
summary(school.std.kaz.merge$ESCS)                    # -5.1 to 3.999; this is the 'vector' (variable) of interest
length(unique(school.std.kaz.merge$CNTSCHID))         # 616 unique school IDS, this is the 'index'
school.ESCS.means <- tapply(school.std.kaz.merge$ESCS, school.std.kaz.merge$CNTSCHID, FUN=function(x)mean(x, na.rm = T))
print(school.ESCS.means)
# Apply logic to PISA
# We get school.ESCS.means
print(school.ESCS.means)                              # we have School ESCS from PISA< we worked it out!
PISA.N.Std.Each.School <- unname(table(school.std.kaz.merge$CNTSCHID))
AVG.PISA.ESCS <- rep(school.ESCS.means, PISA.N.Std.Each.School)
length(AVG.PISA.ESCS)    #19507
school.std.kaz.merge <- cbind.data.frame(school.std.kaz.merge, AVG.PISA.ESCS)

##########check it out### SC001Q01TA #school location
#############All STUDENT-LEVEL VARIABLES################################
seventy.columns <- colnames(school.std.kaz.merge) %in% c("ESCS","ST097Q01TA",
"ST097Q02TA","ST097Q03TA","ST097Q04TA","ST097Q05TA","ST100Q01TA","ST100Q02TA","ST100Q03TA",
"ST100Q04TA","ST211Q01HA","ST211Q02HA","ST211Q03HA","ST213Q01HA","ST213Q02HA","ST213Q03HA","ST213Q04HA",
"ST034Q01TA","ST034Q04TA","ST034Q06TA","ST205Q01HA","ST205Q02HA","ST205Q03HA","ST205Q04HA",
"ST038Q04NA","ST038Q05NA","ST038Q06NA","ST038Q07NA","ST206Q01HA","ST206Q03HA","ST206Q04HA",
```

```
"ST062Q01TA","ST062Q02TA","ST062Q03TA","ST177Q01HA","ST177Q02HA","ST177Q03HA","ST123Q02NA",
"ST123Q03NA","ST123Q04NA","ST207Q01HA", "ST207Q02HA", "ST207Q03HA","ST207Q04HA", "ST207Q05HA",
"EDUSHORT","ST004D01T","SCHSIZE","CNTSCHID", "SCHLTYPE","STAFFSHORT", "CREACTIV",
"SC011","W_FSTUWT","AVG.PISA.ESCS","SC001Q01TA", "PV1MATH","PV2MATH", "PV3MATH","PV4MATH","PV5MATH",
"PV1READ", "PV2READ", "PV3READ", "PV4READ", "PV5READ", "PV1SCIE", "PV2SCIE","PV3SCIE", "PV4SCIE", "PV5SCIE")
print(seventy.columns)
sum(seventy.columns)
seventy.col.df <- school.std.kaz.merge[ , seventy.columns]
####### SUM NA #################
# Identify % missing for each variable
p <- function(x){sum(is.na(x))/length(x)*100}
apply(seventy.col.df, 2, p)
# Use base apply function to check missing data
percent.missing <- apply(seventy.col.df, 2, FUN=function(x)sum(is.na(x))/ length(x)*100)
percent.missing<- round(percent.missing, 2)
summary(percent.missing)
# Remove ST177Q01HA ST177Q02HA ST177Q03HA (meaning, ...)
sixty.7.logical <- !colnames(seventy.col.df) %in%  c("ST177Q01HA","ST177Q02HA","ST177Q03HA")
sixty.7.df <- seventy.col.df[ ,sixty.7.logical]
dim(sixty.7.df)   #19507    67
# Check number of complete cases
total.complete.cases <- sum(complete.cases(sixty.7.df))
percentage.complete <- total.complete.cases/ nrow(sixty.7.df) *100
print(percentage.complete)  #64.87415
# Remove Kazakh students with missing data
#Use logical to remove missing data
removed1 <- sixty.7.df[rowSums(is.na(sixty.7.df)) == 0, ]       # Apply rowSums & is.na
dim(removed1)      # 12655    67
19507-12655        # 6852 students with missing values were removed.
#### Our analysis needs variation in schools (so, we need to remove some schools) # needs school ID
# Remove singleton Schools
table(removed1$CNTSCHID)                                        # Frequency of instances of each school
```

```
sort(table(removed1$CNTSCHID))
singleton.school.IDs <- names(which(sort(table(removed1$CNTSCHID)) == 1))      # Identifying singleton school IDs
length(singleton.school.IDs)                                   # 21 Schools with singletons
non.singletons.Sch.IDs <- !removed1$CNTSCHID %in% singleton.school.IDs        # Creating logical of non-singleton school IDs
removed1 <- removed1[non.singletons.Sch.IDs, ]                             # Retain only non-singletons
dim(removed1) # 12634    67                                      # Check revised dimensions of data frame
##### Checking for within school variation in each variable... (necessary for multilevel modelling with lme4 package)
removed1$ST004D01T
# Student variable "ST004D01T"      (Student (Standardized) Gender)   1= Female ,   2= male
summary(removed1$ST004D01T)       # min 1  , max 2
table(removed1$ST004D01T)                                    #1=6367       #2= 6267
ST004D01T.SDs <- tapply(removed1$ST004D01T, removed1$CNTSCHID, FUN = function(x)sd(x, na.rm=T))
sum(is.na(ST004D01T.SDs))         # Ok, zero
sort(ST004D01T.SDs)               # many single sex schools need to be removed!
schools.no.G.sd <- names(which(sort(ST004D01T.SDs) == 0))
length(schools.no.G.sd)           # 29 schools no gender
SC.ID.gender.v.logical <-  !removed1$CNTSCHID %in% schools.no.G.sd
removed1 <- removed1[SC.ID.gender.v.logical,  ]
dim(removed1)                    # 12432    67
12634-12432 # removed 202 students from 29 schools with single gender.
# Disruptive Student Academic Behavior1
############
removed1$ST097Q01TA     # How often during <test language lessons>: Students don't listen to what the teacher says.
# Student variable ST097Q01TA (Frequency of not listening to what teacher says)
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST097Q01TA<- car::recode(removed1$ST097Q01TA,"
                          1 = 4;
                          2 = 3;
                          3 = 2;
                          4 = 1")
summary(removed1$ST097Q01TA)  # OK min 1 , max 4
table(removed1$ST097Q01TA)    # OK 1,2,3,4
```

```
ST097Q01TA.SDs <- tapply(removed1$ST097Q01TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST097Q01TA.SDs) #
sum(is.na(ST097Q01TA.SDs))  # OK, zero (this check OK but maybe redundant, because all missing data removed anyway)
schools.no.97Q1.sd <- names(which(sort(ST097Q01TA.SDs) == 0))
length(schools.no.97Q1.sd)     # 21 schools have no variation
SC.ID.97Q1.v.logical <-  !removed1$CNTSCHID %in% schools.no.97Q1.sd
removed1 <- removed1[SC.ID.97Q1.v.logical,  ]
dim(removed1)         # 12319    67
12432-12319              # removed 113 students from  21 schools with single answer
removed1$ST097Q02TA
# How often during <test language lessons>: There is noise and disorder.
# Student variable ST097Q02TA (Frequency of noise and disorder)
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST097Q02TA<- car::recode(removed1$ST097Q02TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST097Q02TA)  # OK min 1 , max 4
table(removed1$ST097Q02TA)    # OK 1,2,3,4
ST097Q02TA.SDs <- tapply(removed1$ST097Q02TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST097Q02TA.SDs) #
schools.no.97Q2.sd<- names(which(sort(ST097Q02TA.SDs) == 0))
length(schools.no.97Q2.sd)      # 21 schools
SC.ID.97Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.97Q2.sd
removed1 <- removed1[SC.ID.97Q2.v.logical,  ]
dim(removed1)        # 12203    67
12319-12203         # removed 116 students from 21 schools with single answer
sum(is.na(ST097Q02TA.SDs))  # OK, zero


removed1$ST097Q03TA
```

```
# Student variable ST097Q03TA (Frequency of noise and disorder)
# How often during <test language lessons>: The teacher waits long for students to quiet down.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST097Q03TA<- car::recode(removed1$ST097Q03TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST097Q03TA)  # OK min 1 , max 4
table(removed1$ST097Q03TA)    # OK 1,2,3,4
ST097Q03TA.SDs <- tapply(removed1$ST097Q03TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST097Q03TA.SDs) #
schools.no.97Q3.sd<- names(which(sort(ST097Q03TA.SDs) == 0))
SC.ID.97Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.97Q3.sd
removed1 <- removed1[SC.ID.97Q3.v.logical,  ]
length(schools.no.97Q3.sd)            #11 schools
dim(removed1)       # 12135    67
12203-12135         # removed 68 students from 11 schools with single answer
sum(is.na(ST097Q03TA.SDs))  # OK, zero
removed1$ST097Q04TA
# Student variable ST097Q04TA (Frequency of noise and disorder)
# How often during <test language lessons>: Students cannot work well.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST097Q04TA<- car::recode(removed1$ST097Q04TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST097Q04TA)  # OK min 1 , max 4
table(removed1$ST097Q04TA)    # OK 1,2,3,4
ST097Q04TA.SDs <- tapply(removed1$ST097Q04TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST097Q04TA.SDs) #
```

```
schools.no.97Q4.sd<- names(which(sort(ST097Q04TA.SDs) == 0))
SC.ID.97Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.97Q4.sd
removed1 <- removed1[SC.ID.97Q4.v.logical,  ]
length(schools.no.97Q4.sd)      # 9 schools
dim(removed1)        # 12039    67
12135-12039          # removed 96 students from 9 schools with single answer
sum(is.na(ST097Q04TA.SDs))  # OK, zero
removed1$ST097Q05TA
# Student variable ST097Q05TA (Frequency of noise and disorder)
# How often during <test language lessons>: Students don't start working for a long time after the lesson begins.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST097Q05TA<- car::recode(removed1$ST097Q05TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST097Q05TA)  # OK min 1 , max 4
table(removed1$ST097Q05TA)    # OK 1,2,3,4
ST097Q05TA.SDs <- tapply(removed1$ST097Q05TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST097Q05TA.SDs) #
schools.no.97Q5.sd<- names(which(sort(ST097Q05TA.SDs) == 0))
SC.ID.97Q5.v.logical <- !removed1$CNTSCHID %in% schools.no.97Q5.sd
removed1 <- removed1[SC.ID.97Q5.v.logical,  ]
length(schools.no.97Q5.sd)   # 9 schools
dim(removed1)        # 11980    67
12039-11980          # removed 59 students from 9 schools with single answer
sum(is.na(ST097Q05TA.SDs))  # OK, zero
removed1$ST100Q01TA
# Student variable ST100Q01TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher shows an interest in every student's learning.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST100Q01TA<- car::recode(removed1$ST100Q01TA,"
```

```
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")

summary(removed1$ST100Q01TA)  # OK min 1 , max 4
table(removed1$ST100Q01TA)    # OK 1,2,3,4
ST100Q01TA.SDs <- tapply(removed1$ST100Q01TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST100Q01TA.SDs) #
schools.no.100Q1.sd<- names(which(sort(ST100Q01TA.SDs) == 0))
SC.ID.100Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.100Q1.sd
removed1 <- removed1[SC.ID.100Q1.v.logical,  ]
length(schools.no.100Q1.sd)    # 1 school
dim(removed1)       # 11976    67
11980-11976         # removed 4 students from 1 school with single answer
sum(is.na(ST100Q01TA.SDs))  # OK, zero
removed1$ST100Q02TA
# Student variable ST100Q02TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher gives extra help when students need it
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST100Q02TA<- car::recode(removed1$ST100Q02TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST100Q02TA)  # OK min 1 , max 4
table(removed1$ST100Q02TA)    # OK 1,2,3,4
ST100Q02TA.SDs <- tapply(removed1$ST100Q02TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST100Q02TA.SDs) #
schools.no.100Q2.sd<- names(which(sort(ST100Q02TA.SDs) == 0))
SC.ID.100Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.100Q2.sd
removed1 <- removed1[SC.ID.100Q2.v.logical,  ]
```

```
length(schools.no.100Q2.sd)   #0 schools
dim(removed1)        # 11976    67
11976-11976          # removed zero students from schools with single answer
sum(is.na(ST100Q02TA.SDs))  # OK, zero
removed1$ST100Q03TA
# Student variable ST100Q03TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher helps students with their learning.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever

removed1$ST100Q03TA<- car::recode(removed1$ST100Q03TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST100Q03TA)  # OK min 1 , max 4
table(removed1$ST100Q03TA)    # OK 1,2,3,4
ST100Q03TA.SDs <- tapply(removed1$ST100Q03TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST100Q03TA.SDs) #
schools.no.100Q3.sd<- names(which(sort(ST100Q03TA.SDs) == 0))
SC.ID.100Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.100Q3.sd
removed1 <- removed1[SC.ID.100Q3.v.logical,  ]
length(schools.no.100Q3.sd)   # 1 school
dim(removed1)        #  11973    67
11976-11973          # removed 3 students from  1 school with single answer
sum(is.na(ST100Q03TA.SDs))  # OK, zero
removed1$ST100Q04TA
# Student variable ST100Q04TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher continues teaching until the students understands.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
removed1$ST100Q04TA<- car::recode(removed1$ST100Q04TA,"
                              1 = 4;
                              2 = 3;
```

```
                              3 = 2;
                              4 = 1")
summary(removed1$ST100Q04TA)  # OK min 1 , max 4
table(removed1$ST100Q04TA)    # OK 1,2,3,4
ST100Q04TA.SDs <- tapply(removed1$ST100Q04TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST100Q04TA.SDs) #
schools.no.100Q4.sd<- names(which(sort(ST100Q04TA.SDs) == 0))
SC.ID.100Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.100Q4.sd
removed1 <- removed1[SC.ID.100Q4.v.logical,  ]
length(schools.no.100Q4.sd)   # 1 school
dim(removed1)        # 11967    67
11973-11967          # removed 6 students from 1 school with single answer
sum(is.na(ST100Q04TA.SDs))  # OK, zero
removed1$ST211Q01HA
# Student variable ST211Q01HA (Agreement: The teacher made me feel confident in my ability to do well in the course..)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST211Q01HA)  # OK min 1 , max 4
table(removed1$ST211Q01HA)    # OK 1,2,3,4
ST211Q01HA.SDs <- tapply(removed1$ST211Q01HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST211Q01HA.SDs) #
schools.no.211Q1.sd<- names(which(sort(ST211Q01HA.SDs) == 0))
SC.ID.211Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.211Q1.sd
removed1 <- removed1[SC.ID.211Q1.v.logical,  ]
length(schools.no.211Q1.sd)    # 0 school
dim(removed1)        # 11967    67
11967-11967         # removed 0 students from  0 school with single answer
sum(is.na(ST211Q01HA.SDs))  # OK, zero


removed1$ST211Q02HA
# Student variable ST211Q02HA (Agreement: Thinking of past two <test language lessons>:
#                              The teacher listened to my view on how to do things.)
```

```
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST211Q02HA)  # OK min 1 , max 4
table(removed1$ST211Q02HA)    # OK 1,2,3,4
ST211Q02HA.SDs <- tapply(removed1$ST211Q02HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST211Q02HA.SDs) #
schools.no.211Q2.sd<- names(which(sort(ST211Q02HA.SDs) == 0))
SC.ID.211Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.211Q2.sd
removed1 <- removed1[SC.ID.211Q2.v.logical,  ]
length(schools.no.211Q2.sd) # 0 schools
dim(removed1)        # 11967    67
11967-11967          # removed zero students from schools with single answer
sum(is.na(ST211Q02HA.SDs))  # OK, zero
removed1$ST211Q03HA
# Student variable ST211Q03HA (Agreement: Thinking of past two <test language lessons>: I felt that my teacher understood me. )
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST211Q03HA)  # OK min 1 , max 4
table(removed1$ST211Q03HA)    # OK 1,2,3,4
ST211Q03HA.SDs <- tapply(removed1$ST211Q03HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST211Q03HA.SDs) #
schools.no.211Q3.sd<- names(which(sort(ST211Q03HA.SDs) == 0))
SC.ID.211Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.211Q3.sd
removed1 <- removed1[SC.ID.211Q3.v.logical,  ]
length(schools.no.211Q3.sd) # 0 school
dim(removed1)        # 11967    67
11967-11967          # removed 0 students from 0 school with single answer
sum(is.na(ST211Q03HA.SDs))  # OK, zero
removed1$ST213Q01HA
# Student variable ST213Q01HA ( Agreement:  It was clear to me that the teacher liked teaching us.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST213Q01HA)  # OK min 1 , max 4
table(removed1$ST213Q01HA)    # OK 1,2,3,4
ST213Q01HA.SDs <- tapply(removed1$ST213Q01HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
```

```
sort(ST213Q01HA.SDs) #
schools.no.213Q1.sd<- names(which(sort(ST213Q01HA.SDs) == 0))
SC.ID.213Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.213Q1.sd
removed1 <- removed1[SC.ID.213Q1.v.logical,  ]
length(schools.no.213Q1.sd) # 2 schools
dim(removed1)       # 11953    67
11967-11953         # removed 14 students from 2 schools with single answer
sum(is.na(ST213Q01HA.SDs))  # OK, zero
removed1$ST213Q02HA
# Student variable ST213Q02HA ( Agreement:  The enthusiasm of the teacher inspired me.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST213Q02HA)  # OK min 1 , max 4
table(removed1$ST213Q02HA)    # OK 1,2,3,4
ST213Q02HA.SDs <- tapply(removed1$ST213Q02HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST213Q02HA.SDs) #
schools.no.213Q2.sd<- names(which(sort(ST213Q02HA.SDs) == 0))
SC.ID.213Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.213Q2.sd
removed1 <- removed1[SC.ID.213Q2.v.logical,  ]
length(schools.no.213Q2.sd) # 0 schools
dim(removed1)       # 11953    67
11953-11953         # removed zero students from schools with single answer
sum(is.na(ST213Q02HA.SDs))  # OK, zero
removed1$ST213Q03HA
# Student variable ST213Q03HA ( Agreement:  It was clear that the teacher likes to deal with the topic of the lesson.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST213Q03HA)  # OK min 1 , max 4
table(removed1$ST213Q03HA)    # OK 1,2,3,4
ST213Q03HA.SDs <- tapply(removed1$ST213Q03HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST213Q03HA.SDs) #
schools.no.213Q3.sd<- names(which(sort(ST213Q03HA.SDs) == 0))
SC.ID.213Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.213Q3.sd
removed1 <- removed1[SC.ID.213Q3.v.logical,  ]
```

```
length(schools.no.213Q3.sd) # 1 school
dim(removed1)       # 11947    67
11953-11947         # removed 6 students from  1 school with single answer
sum(is.na(ST213Q03HA.SDs))  # OK, zero
removed1$ST213Q04HA
# Student variable ST213Q04HA ( Agreement:  The teacher showed enjoyment in teaching.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST213Q04HA)  # OK min 1 , max 4
table(removed1$ST213Q04HA)    # OK 1,2,3,4
ST213Q04HA.SDs <- tapply(removed1$ST213Q04HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST213Q04HA.SDs) #
schools.no.213Q4.sd<- names(which(sort(ST213Q04HA.SDs) == 0))
SC.ID.213Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.213Q4.sd
removed1 <- removed1[SC.ID.213Q4.v.logical,  ]
length(schools.no.213Q4.sd) # 1 school
dim(removed1)       # 11935    67
11947-11935          # removed 12 students from 1 school with single answer
sum(is.na(ST213Q04HA.SDs))  # OK, zero
removed1$ST034Q01TA
# Student variable ST034Q01TA ( Agreement:   I feel like an outsider (or left out of things) at school.)
# 1=Strongly agree , 2=Agree, 3=Disagree  4=Strongly disagree
removed1$ST034Q01TA<- car::recode(removed1$ST034Q01TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST034Q01TA)  # OK min 1 , max 4
table(removed1$ST034Q01TA)    # OK 1,2,3,4
ST034Q01TA.SDs <- tapply(removed1$ST034Q01TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST034Q01TA.SDs) #
schools.no.34Q1.sd<- names(which(sort(ST034Q01TA.SDs) == 0))
SC.ID.34Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.34Q1.sd
```

```
removed1 <- removed1[SC.ID.34Q1.v.logical,  ]
length(schools.no.34Q1.sd) # 1 school
dim(removed1)        # 11928    67
11935-11928          # removed seven students from 1 school with single answer
sum(is.na(ST034Q01TA.SDs))  # OK, zero


removed1$ST034Q04TA
# Student variable ST034Q01TA ( Agreement:  Thinking about your school: I feel awkward and out of place in my school.)
# 1=Strongly agree , 2=Agree, 3=Disagree  4=Strongly disagree
removed1$ST034Q04TA<- car::recode(removed1$ST034Q04TA,"
                              1 = 4;
                              2 = 3;
                              3 = 2;
                              4 = 1")
summary(removed1$ST034Q04TA)  # OK min 1 , max 4
table(removed1$ST034Q04TA)    # OK 1,2,3,4
ST034Q04TA.SDs <- tapply(removed1$ST034Q04TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST034Q04TA.SDs) #
schools.no.34Q4.sd<- names(which(sort(ST034Q04TA.SDs) == 0))
SC.ID.34Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.34Q4.sd
removed1 <- removed1[SC.ID.34Q4.v.logical,  ]
length(schools.no.34Q4.sd) # 3 schools
dim(removed1)        # 11905    67
11928-11905          # removed 23 students from 3 schools with single answer
sum(is.na(ST034Q04TA.SDs))  # OK, zero
removed1$ST034Q06TA
# Student variable ST036Q01TA ( Agreement:   Thinking about your school: I feel lonely at school.)
# 1=Strongly agree , 2=Agree, 3=Disagree  4=Strongly disagree
removed1$ST034Q06TA<- car::recode(removed1$ST034Q06TA,"
                              1 = 4;
                              2 = 3;
```

```
                                3 = 2;
                                4 = 1")
summary(removed1$ST034Q06TA)  # OK min 1 , max 4
table(removed1$ST034Q06TA)    # OK 1,2,3,4
ST034Q06TA.SDs <- tapply(removed1$ST034Q06TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST034Q06TA.SDs) #
schools.no.34Q6.sd<- names(which(sort(ST034Q06TA.SDs) == 0))
SC.ID.34Q6.v.logical <- !removed1$CNTSCHID %in% schools.no.34Q6.sd
removed1 <- removed1[SC.ID.34Q6.v.logical,  ]
length(schools.no.34Q6.sd) # 0 school
dim(removed1)        # 11905    67
11905-11905          # removed zero students from schools with single answer
sum(is.na(ST034Q06TA.SDs))  # OK, zero
removed1$ST123Q02NA
# Student variable ST123Q02NA ( Agreement:   My parents support my educational efforts and achievements.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST123Q02NA)  # OK min 1 , max 4
table(removed1$ST123Q02NA)    # OK 1,2,3,4
ST123Q02NA.SDs <- tapply(removed1$ST123Q02NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST123Q02NA.SDs) #
schools.no.123Q2.sd<- names(which(sort(ST123Q02NA.SDs) == 0))
SC.ID.123Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.123Q2.sd
removed1 <- removed1[SC.ID.123Q2.v.logical,  ]
length(schools.no.123Q2.sd) # 0 school
dim(removed1)        # 11905    67
11905-11905          # removed zero students from schools with single answer
sum(is.na(ST123Q02NA.SDs))  # OK, zero
removed1$ST123Q03NA
# Student variable ST123Q03NA ( Agreement:  My parents support me when I am facing difficulties at school.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST123Q03NA)  # OK min 1 , max 4
table(removed1$ST123Q03NA)    # OK 1,2,3,4
```

```
ST123Q03NA.SDs <- tapply(removed1$ST123Q03NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST123Q03NA.SDs) #
schools.no.123Q3.sd<- names(which(sort(ST123Q03NA.SDs) == 0))
SC.ID.123Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.123Q3.sd
removed1 <- removed1[SC.ID.123Q3.v.logical,  ]
length(schools.no.123Q3.sd) # 0 school
dim(removed1)        # 11905    67
11905-11905          # removed zero students from schools with single answer
sum(is.na(ST123Q03NA.SDs))  # OK, zero
removed1$ST123Q04NA
# Student variable ST123Q04NA ( Agreement:   My parents encourage me to be confident.)
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
summary(removed1$ST123Q04NA)  # OK min 1 , max 4
table(removed1$ST123Q04NA)    # OK 1,2,3,4
ST123Q04NA.SDs <- tapply(removed1$ST123Q04NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST123Q04NA.SDs) #
schools.no.123Q4.sd<- names(which(sort(ST123Q04NA.SDs) == 0))
SC.ID.123Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.123Q4.sd
removed1 <- removed1[SC.ID.123Q4.v.logical,  ]
length(schools.no.123Q4.sd) # 0 school
dim(removed1)        # 11905    67
11905-11905         # removed zero students from schools with single answer
sum(is.na(ST123Q04NA.SDs))  # OK, zero
removed1$ST205Q01HA
# Student variable ST205Q01HA ( How true: Students seem to value competition.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST205Q01HA)  # OK min 1 , max 4
table(removed1$ST205Q01HA)    # OK 1,2,3,4
ST205Q01HA.SDs <- tapply(removed1$ST205Q01HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST205Q01HA.SDs) #
schools.no.205Q1.sd<- names(which(sort(ST205Q01HA.SDs) == 0))
SC.ID.205Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.205Q1.sd
```

```
removed1 <- removed1[SC.ID.205Q1.v.logical,  ]
length(schools.no.205Q1.sd) # 0 school
dim(removed1)       # 11905    67
11905-11905         # removed zero students from schools with single answer
sum(is.na(ST205Q01HA.SDs))  # OK, zero
removed1$ST205Q02HA
# Student variable ST205Q02HA ( How true: It seems that students are competing with each other.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST205Q02HA)  # OK min 1 , max 4
table(removed1$ST205Q02HA)    # OK 1,2,3,4
ST205Q02HA.SDs <- tapply(removed1$ST205Q02HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST205Q02HA.SDs) #
schools.no.205Q2.sd<- names(which(sort(ST205Q02HA.SDs) == 0))
SC.ID.205Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.205Q2.sd
removed1 <- removed1[SC.ID.205Q2.v.logical,  ]
length(schools.no.205Q2.sd) # 1 school
dim(removed1)       # 11899    67
11905-11899         # removed 6 students from 1 school with single answer
sum(is.na(ST205Q02HA.SDs))  # OK, zero
removed1$ST205Q03HA
# Student variable ST205Q03HA ( How true:  Students seem to share the feeling that competing with each other is important.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST205Q03HA)  # OK min 1 , max 4
table(removed1$ST205Q03HA)    # OK 1,2,3,4
ST205Q03HA.SDs <- tapply(removed1$ST205Q03HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST205Q03HA.SDs) #
schools.no.205Q3.sd<- names(which(sort(ST205Q03HA.SDs) == 0))
SC.ID.205Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.205Q3.sd
removed1 <- removed1[SC.ID.205Q3.v.logical,  ]
length(schools.no.205Q3.sd) # 0 school
dim(removed1)       # 11899    67
11899-11899         # removed zero students from schools with single answer
```

```
sum(is.na(ST205Q03HA.SDs))  # OK, zero
removed1$ST205Q04HA
# Student variable ST205Q04HA ( How true: Students feel that they are being compared with others.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST205Q04HA)  # OK min 1 , max 4
table(removed1$ST205Q04HA)    # OK 1,2,3,4
ST205Q04HA.SDs <- tapply(removed1$ST205Q04HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST205Q04HA.SDs) #
schools.no.205Q4.sd<- names(which(sort(ST205Q04HA.SDs) == 0))
SC.ID.205Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.205Q4.sd
removed1 <- removed1[SC.ID.205Q4.v.logical,  ]
length(schools.no.205Q4.sd) # 0 school
dim(removed1)        # 11899    67
11899-11899          # removed zero students from schools with single answer
sum(is.na(ST205Q04HA.SDs))  # OK, zero
removed1$ST062Q01TA
# Student variable ST062Q01TA (  How often: I <skipped> a whole school day.)
# 1=Never , 2=One or two times, 3=Three or four times,  4=Five or more times
summary(removed1$ST062Q01TA)  # OK min 1 , max 4
table(removed1$ST062Q01TA)    # OK 1,2,3,4
ST062Q01TA.SDs <- tapply(removed1$ST062Q01TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST062Q01TA.SDs) #
schools.no.62Q1.sd<- names(which(sort(ST062Q01TA.SDs) == 0))
SC.ID.62Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.62Q1.sd
removed1 <- removed1[SC.ID.62Q1.v.logical,  ]
length(schools.no.62Q1.sd) # 0 school
dim(removed1)        # 11899    67
11899-11899          # removed 0 students from schools with single answer
sum(is.na(ST062Q01TA.SDs))  # OK, zero
removed1$ST062Q02TA
# Student variable ST062Q02TA (  How often:  I <skipped> some classes.)
# 1=Never , 2=One or two times, 3=Three or four times,  4=Five or more times
```

```
summary(removed1$ST062Q02TA)  # OK min 1 , max 4
table(removed1$ST062Q02TA)    # OK 1,2,3,4
ST062Q02TA.SDs <- tapply(removed1$ST062Q02TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST062Q02TA.SDs) #
schools.no.62Q2.sd<- names(which(sort(ST062Q02TA.SDs) == 0))
SC.ID.62Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.62Q2.sd
removed1 <- removed1[SC.ID.62Q2.v.logical,  ]
length(schools.no.62Q2.sd) # 0 school
dim(removed1)        # 11899    67
11899-11899          # removed zero students from schools with single answer
sum(is.na(ST062Q02TA.SDs))  # OK, zero
removed1$ST062Q03TA
# Student variable ST062Q03TA (  How often:  I arrived late for school)
# 1=Never , 2=One or two times, 3=Three or four times,  4=Five or more times
summary(removed1$ST062Q03TA)  # OK min 1 , max 4
table(removed1$ST062Q03TA)    # OK 1,2,3,4
ST062Q03TA.SDs <- tapply(removed1$ST062Q03TA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST062Q03TA.SDs) #
schools.no.62Q3.sd<- names(which(sort(ST062Q03TA.SDs) == 0))
SC.ID.62Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.62Q3.sd
removed1 <- removed1[SC.ID.62Q3.v.logical,  ]
length(schools.no.62Q3.sd) # 0 school
dim(removed1)        # 11899    67
11899-11899          # removed zero students from schools with single answer
sum(is.na(ST062Q03TA.SDs))  # OK, zero
removed1$ST038Q04NA
# Student variable ST038Q04NA (   How often: Other students made fun of me.)
# 1=Never or almost never , 2=A few times a year, 3=A few times a month,  4=Once a week or more
summary(removed1$ST038Q04NA)  # OK min 1 , max 4
table(removed1$ST038Q04NA)    # OK 1,2,3,4
ST038Q04NA.SDs <- tapply(removed1$ST038Q04NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST038Q04NA.SDs) #
```

```
schools.no.38Q4.sd<- names(which(sort(ST038Q04NA.SDs) == 0))
SC.ID.38Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.38Q4.sd
removed1 <- removed1[SC.ID.38Q4.v.logical,  ]
length(schools.no.38Q4.sd) # 12 school
dim(removed1)        # 11738    67
11899-11738          # removed 161 students from 12 schools with single answer
sum(is.na(ST038Q04NA.SDs))  # OK, zero
removed1$ST038Q05NA
# Student variable ST038Q05NA (   How often:  I was threatened by other students.)
# 1=Never or almost never , 2=A few times a year, 3=A few times a month,  4=Once a week or more
summary(removed1$ST038Q05NA)  # OK min 1 , max 4
table(removed1$ST038Q05NA)    # OK 1,2,3,4
ST038Q05NA.SDs <- tapply(removed1$ST038Q05NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST038Q05NA.SDs) #
schools.no.38Q5.sd<- names(which(sort(ST038Q05NA.SDs) == 0))
SC.ID.38Q5.v.logical <- !removed1$CNTSCHID %in% schools.no.38Q5.sd
removed1 <- removed1[SC.ID.38Q5.v.logical,  ]
length(schools.no.38Q5.sd) # 10 school
dim(removed1)        # 11596    67
11738-11596          # removed 142 students from 10 schools with single answer
sum(is.na(ST038Q05NA.SDs))  # OK, zero


removed1$ST038Q06NA
# Student variable ST038Q06NA (   How often: Other students took away or destroyed things that belonged to me.)
# 1=Never or almost never , 2=A few times a year, 3=A few times a month,  4=Once a week or more
summary(removed1$ST038Q06NA)  # OK min 1 , max 4
table(removed1$ST038Q06NA)    # OK 1,2,3,4
ST038Q06NA.SDs <- tapply(removed1$ST038Q06NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST038Q06NA.SDs) #
schools.no.38Q6.sd<- names(which(sort(ST038Q06NA.SDs) == 0))
SC.ID.38Q6.v.logical <- !removed1$CNTSCHID %in% schools.no.38Q6.sd
```

```
removed1 <- removed1[SC.ID.38Q6.v.logical,  ]
length(schools.no.38Q6.sd) # 1 school
dim(removed1)       # 11572    67
11596-11572        # removed 24 students from 1 school with single answer
sum(is.na(ST038Q06NA.SDs))  # OK, zero
removed1$ST038Q07NA
# Student variable ST038Q07NA (    How often: I got hit or pushed around by other students.)
# 1=Never or almost never , 2=A few times a year, 3=A few times a month,  4=Once a week or more
summary(removed1$ST038Q07NA)  # OK min 1 , max 4
table(removed1$ST038Q07NA)    # OK 1,2,3,4
ST038Q07NA.SDs <- tapply(removed1$ST038Q07NA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST038Q07NA.SDs) #
schools.no.38Q7.sd<- names(which(sort(ST038Q07NA.SDs) == 0))
SC.ID.38Q7.v.logical <- !removed1$CNTSCHID %in% schools.no.38Q7.sd
removed1 <- removed1[SC.ID.38Q7.v.logical,  ]
length(schools.no.38Q7.sd) # 1 school
dim(removed1)       # 11535    67
11572-11535        # removed 37 students from 1 school with single answer
sum(is.na(ST038Q07NA.SDs))  # OK, zero
removed1$ST206Q01HA
# Student variable ST206Q01HA (    How true: Students seem to value cooperation.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST206Q01HA)  # OK min 1 , max 4
table(removed1$ST206Q01HA)    # OK 1,2,3,4
ST206Q01HA.SDs <- tapply(removed1$ST206Q01HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST206Q01HA.SDs) #
schools.no.206Q1.sd<- names(which(sort(ST206Q01HA.SDs) == 0))
SC.ID.206Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.206Q1.sd
removed1 <- removed1[SC.ID.206Q1.v.logical,  ]
length(schools.no.206Q1.sd) # 0 school
dim(removed1)       # 11535    67
11535-11535        # removed zero students from schools with single answer
```

```
sum(is.na(ST206Q01HA.SDs))  # OK, zero
removed1$ST206Q03HA
# Student variable ST206Q03HA (   How true: Students seem to share the feeling that cooperating with each other is important.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST206Q03HA)  # OK min 1 , max 4
table(removed1$ST206Q03HA)    # OK 1,2,3,4
ST206Q03HA.SDs <- tapply(removed1$ST206Q03HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST206Q03HA.SDs) #
schools.no.206Q3.sd<- names(which(sort(ST206Q03HA.SDs) == 0))
SC.ID.206Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.206Q3.sd
removed1 <- removed1[SC.ID.206Q3.v.logical,  ]
length(schools.no.206Q3.sd) # 0 school
dim(removed1)       # 11535    67
11535-11535         # removed zero students from schools with single answer
sum(is.na(ST206Q03HA.SDs))  # OK, zero
removed1$ST206Q04HA
# Student variable ST206Q04HA (   How true: Students feel that they are encouraged to cooperate with others.)
# 1=Not at all true , 2=Slightly true, 3=Very true  4=Extremely true
summary(removed1$ST206Q04HA)  # OK min 1 , max 4
table(removed1$ST206Q04HA)    # OK 1,2,3,4
ST206Q04HA.SDs <- tapply(removed1$ST206Q04HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST206Q04HA.SDs) #
schools.no.206Q4.sd<- names(which(sort(ST206Q04HA.SDs) == 0))
SC.ID.206Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.206Q4.sd
removed1 <- removed1[SC.ID.206Q4.v.logical,  ]
length(schools.no.206Q4.sd) # 0 school
dim(removed1)       # 11535    67
11535-11535         # removed zero students from schools with single answer
sum(is.na(ST206Q04HA.SDs))  # OK, zero
### Anti-bullying attitude
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
#207                  (Agree: It irritates me when nobody defends bullied students.)
```

```
removed1$ST207Q01HA
summary(removed1$ST207Q01HA)      # OK min 1 , max 4
table(removed1$ST207Q01HA)        # OK 1,2,3,4
ST207Q01HA.SDs <- tapply(removed1$ST207Q01HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST207Q01HA.SDs) #
schools.no.207Q1.sd<- names(which(sort(ST207Q01HA.SDs) == 0))
SC.ID.207Q1.v.logical <- !removed1$CNTSCHID %in% schools.no.207Q1.sd
removed1 <- removed1[SC.ID.207Q1.v.logical,  ]
length(schools.no.207Q1.sd) # 0 school
dim(removed1)        # 11535    67
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
#207                   (Agree: It is a good thing to help students who can't defend themselves.)
removed1$ST207Q02HA
summary(removed1$ST207Q02HA)      # OK min 1 , max 4
table(removed1$ST207Q02HA)        # OK 1,2,3,4
ST207Q02HA.SDs <- tapply(removed1$ST207Q02HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST207Q02HA.SDs) #
schools.no.207Q2.sd<- names(which(sort(ST207Q02HA.SDs) == 0))
SC.ID.207Q2.v.logical <- !removed1$CNTSCHID %in% schools.no.207Q2.sd
removed1 <- removed1[SC.ID.207Q2.v.logical,  ]
length(schools.no.207Q2.sd) # 1 school
dim(removed1)        # 11528    67
11535-11528          #removed 7 students from 1 school
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
#207                   (Agree: It is a wrong thing to join in bullying..)
removed1$ST207Q03HA
summary(removed1$ST207Q03HA)      # OK min 1 , max 4
table(removed1$ST207Q03HA)        # OK 1,2,3,4
ST207Q03HA.SDs <- tapply(removed1$ST207Q03HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST207Q03HA.SDs) #
schools.no.207Q3.sd<- names(which(sort(ST207Q03HA.SDs) == 0))
SC.ID.207Q3.v.logical <- !removed1$CNTSCHID %in% schools.no.207Q3.sd
```

```
removed1 <- removed1[SC.ID.207Q3.v.logical,  ]
length(schools.no.207Q3.sd) # 0 school
dim(removed1)       # 11528    67
# removed 0 students
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
#207                  (Agree: I feel bad seeing other students bullied.)
removed1$ST207Q04HA
summary(removed1$ST207Q04HA)      # OK min 1 , max 4
table(removed1$ST207Q04HA)       # OK 1,2,3,4
ST207Q04HA.SDs <- tapply(removed1$ST207Q04HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST207Q04HA.SDs) #
schools.no.207Q4.sd<- names(which(sort(ST207Q04HA.SDs) == 0))
SC.ID.207Q4.v.logical <- !removed1$CNTSCHID %in% schools.no.207Q4.sd
removed1 <- removed1[SC.ID.207Q4.v.logical,  ]
length(schools.no.207Q4.sd) # 0 school
dim(removed1)       # 11528    67
# removed 0 students
# 1=Strongly disagree , 2=Disagree, 3=Agree  4=Strongly agree
#207                  (Agree:  I like it when someone stands up for other students who are being bullied.)
removed1$ST207Q05HA
summary(removed1$ST207Q05HA)      # OK min 1 , max 4
table(removed1$ST207Q05HA)       # OK 1,2,3,4
ST207Q05HA.SDs <- tapply(removed1$ST207Q05HA, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ST207Q05HA.SDs) #
schools.no.207Q5.sd<- names(which(sort(ST207Q05HA.SDs) == 0))
SC.ID.207Q5.v.logical <- !removed1$CNTSCHID %in% schools.no.207Q5.sd
removed1 <- removed1[SC.ID.207Q5.v.logical,  ]
length(schools.no.207Q4.sd) # 0 school
dim(removed1)       # 11528    67
# removed 0 students
removed1$ESCS
# Student variable ESCS (  Index of economic, social and cultural status)
```

```
summary(removed1$ESCS)  # OK min -4.5977  , max 3.9991
table(removed1$ESCS)    #
ESCS.SDs <- tapply(removed1$ESCS, removed1$CNTSCHID, FUN=function(x)sd(x, na.rm=T))
sort(ESCS.SDs) #
schools.no.ESCS.sd<- names(which(sort(ESCS.SDs) == 0))
SC.ID.ESCS.v.logical <- !removed1$CNTSCHID %in% schools.no.ESCS.sd
removed1 <- removed1[SC.ID.ESCS.v.logical,  ]
dim(removed1)        # 11528    67
11528-11528          # removed zero students from schools with single answer
sum(is.na(ESCS.SDs))  # OK, zero
length(unique(removed1$CNTSCHID))     # 426 schools
################################################################################
# final data includes 426 schools, 11528 students and 67 school & student variables
# Focus now is checking School variables
dim(removed1)       # 11528    67
# Location
summary(removed1$SC001Q01TA)     #   min 1,   max  5
table(removed1$SC001Q01TA)  # fine, 1(village fewer than 3000), 2, 3, 4, 5(large city over 1 mill),
mean(removed1$SC001Q01TA)   #
removed1$SC001Q01TA <- removed1$SC001Q01TA - mean(removed1$SC001Q01TA)
mean(removed1$SC001Q01TA)
dim(removed1)      # 11528    67
# SCHLTYPE (1 private, 2 partner, 3 public)      #School Ownership
summary(removed1$SCHLTYPE)        # min 1  ,    max 3
table(removed1$SCHLTYPE)          # OK 1,2,3: 561   471   10496
dim(removed1)                     # 11528    67
library("car")
# dummy coding
# Dummy coding the private only variable (priv = 2, non-priv = 1)
priv.only <-car::recode(removed1$SCHLTYPE,
                    "1 = 2;
                    2 = 1;
```

```
                    3 = 1")
table(priv.only)  # expectation: 2 (561   private) 1 (10967   )
561+10967
length(priv.only)
dim(removed1)     #11528    67
# Dummy coding the partner only variable (part = 2, non-partner 1)
part.only <-car::recode(removed1$SCHLTYPE,"
                    1 = 1;
                    2 = 2;
                    3 = 1")
table(part.only)  # expectation 2 (471 partner, 11057 non-partner)
471+11057
# combining new variables into the dataframe
removed1 <- cbind.data.frame(removed1, priv.only, part.only)
dim(removed1)                  # 11528    69
removed1$SCHLTYPE <-NULL
dim(removed1)    # 11528    68
# SCHSIZE                      # School Size (Sum)
summary(removed1$SCHSIZE)          # min -1.5012     ,   max 3.8481
table(removed1$SCHSIZE)            # OK
dim(removed1)                  # 11528    68
removed1$SCHSIZE <- scale(removed1$SCHSIZE)
summary(removed1$SCHSIZE)
removed1$SCHSIZE <- removed1$SCHSIZE - mean(removed1$SCHSIZE)
mean(removed1$SCHSIZE)
# CREACTIV                     # Creative extra-curricular activities (Sum)
summary(removed1$CREACTIV)         # min 0   ,   max 3
table(removed1$CREACTIV)           # OK 0,1,2,3
dim(removed1)                  # 11528    68
# EDUSHORT                     # Teacher's view on educational material shortage (WLE)
summary(removed1$EDUSHORT)         # min -1.42120   ,   max 2.95950
dim(removed1)                  # 11528    68
```

```
vector.mean.zero <- c(1,2,3) - mean(c(1,2,3))
print(vector.mean.zero)
removed1$EDUSHORT <- removed1$EDUSHORT - mean(removed1$EDUSHORT)
mean(removed1$EDUSHORT)

# STAFFSHORT                    # Teacher's view on staff shortage (WLE)
summary(removed1$STAFFSHORT)      # min -1.11296   ,   max 4.50774
dim(removed1)                   # 11528    68
removed1$STAFFSHORT <- removed1$STAFFSHORT - mean(removed1$STAFFSHORT)
mean(removed1$STAFFSHORT)
# AVG.PISA.ESCS
summary(removed1$AVG.PISA.ESCS)    # min -1.3664 , max = 0.7213
dim(removed1)                   # 11528    68
removed1$AVG.PISA.ESCS <- removed1$AVG.PISA.ESCS - mean(removed1$AVG.PISA.ESCS)
options(scipen=999)
mean(removed1$AVG.PISA.ESCS)
dim(removed1)                   # 11528    68
#### Check skewness with apply function ####
variable.skewness <- apply(removed1, 2, FUN=function(x)e1071::skewness(x))
above.abs.2 <- which(variable.skewness > 2 | variable.skewness < -2)
# SCHLTYPE ST038Q05NA ST038Q06NA ST038Q07NA  priv.only  part.only
#    3        44        45        46        53        5
variable.skewness[names(variable.skewness) == names(above.abs.2)]
# ST038Q05NA ST038Q06NA ST038Q07NA  priv.only  part.only
# 2.147684   2.059455   2.197186   4.182383   4.691505
# these three are moderately positively skewed with long tails to the right. Only just above 2 so we can leave as is and not transform.
# priv.only and part.only are binary variables so not important.
# ST038Q05NA      (These 3 are related to bullying)
# ST038Q06NA
# ST038Q07NA
# priv.only
# part.only
```

```
str(removed1)
head(removed1)
apply(removed1, 2, function(x)mean(x))     # average of variable (mean)
dim(removed1)    #11528    68
############################## NULL #################################
# Null model explores proportion of variance attributable to group effects
#### Explore variation in PV1MATH
NULL.MODEL.PV1MATH <- lme4::lmer(PV1MATH ~ 1                                    ## There are no predictors, only group mean is fixed
                        + (1 | CNTSCHID),                       ## Denotes that the variable of interest is nested in schools
                        data = removed1,                        ## Specifies the dataset used in the analysis
                        REML=F,
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV1MATH)                                     ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV1MATH),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.3032414 *Study one-way ANOVA mathematics to understand basic mechanics of this.
#### Explore Basics of one-way ANOVA!
oneway.test(PV1MATH ~ CNTSCHID, data = removed1, var.equal = T )
anova(lm(removed1$PV1MATH ~ removed1$CNTSCHID))
#### Explore variation in PV1READ
NULL.MODEL.PV1 <- lme4::lmer(PV1READ ~ 1                                 ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                       ## Denotes that the variable of interest is nested in schools
                    data = removed1,                        ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV1)                                     ## Need to extract variance components from the null model
```

```
Var.comp <- print(VarCorr(NULL.MODEL.PV1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.3724673
#### Explore variation in PV1SCIE
NULL.MODEL.PV1 <- lme4::lmer(PV1SCIE ~ 1                    ## There are no predictors, only group mean is fixed
                + (1 | CNTSCHID),                ## Denotes that the variable of interest is nested in schools
                data = removed1,                ## Specifies the dataset used in the analysis
                REML=F,
                control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV1)                    ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.3783766
removed1$PV2MATH
#### Explore variation in PV2MATH
NULL.MODEL.PV2MATH <- lme4::lmer(PV2MATH ~ 1                    ## There are no predictors, only group mean is fixed
                + (1 | CNTSCHID),                ## Denotes that the variable of interest is nested in schools
                data = removed1,                ## Specifies the dataset used in the analysis
                REML=F,
                control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV2MATH)                    ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV2MATH),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.3030115
removed1$PV3MATH
#### Explore variation in PV3MATH
NULL.MODEL.PV3MATH <- lme4::lmer(PV3MATH ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                    data = removed1,                  ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV3MATH)                           ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV3MATH),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.3189917
removed1$PV4MATH
#### Explore variation in PV3MATH
NULL.MODEL.PV4MATH <- lme4::lmer(PV4MATH ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                    data = removed1,                  ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV4MATH)                           ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV4MATH),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.3053233
removed1$PV5MATH
#### Explore variation in PV5MATH
NULL.MODEL.PV5MATH <- lme4::lmer(PV5MATH ~ 1                              ## There are no predictors, only group mean is fixed
                         + (1 | CNTSCHID),                        ## Denotes that the variable of interest is nested in schools
                         data = removed1,                         ## Specifies the dataset used in the analysis
                         REML=F,
                         control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                          optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV5MATH)                                   ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV5MATH),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.319588
#### Explore variation in BULLYING
removed1$ST038Q04NA # Other students make fun of me
summary(removed1$ST038Q04NA)
sd(removed1$ST038Q04NA)
NULL.MODEL.PV1 <- lme4::lmer(ST038Q04NA ~ 1                              ## There are no predictors, only group mean is fixed
                       + (1 | CNTSCHID),                          ## Denotes that the variable of interest is nested in schools
                       data = removed1,                           ## Specifies the dataset used in the analysis
                       REML=F,
                       control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                        optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
```

```
summary(NULL.MODEL.PV1)                          ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.01520859 2% variance between schools
## Bullying 2
removed1
summary(removed1$ST038Q04NA)
sd(removed1$ST038Q04NA)
NULL.MODEL.PV1 <- lme4::lmer(ST038Q04NA ~ 1                   ## There are no predictors, only group mean is fixed
                  + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                  data = removed1,                  ## Specifies the dataset used in the analysis
                  REML=F,
                  control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV1)                          ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    #  0.01520859
#### Explore variation in TEACHER DEDICAATION
removed1$ST213Q01HA # It was clear to me that the teacher liked teaching us.
summary(removed1$ST213Q01HA)
sd(removed1$ST213Q01HA)
NULL.MODEL.PV1 <- lme4::lmer(ST213Q01HA ~ 1                   ## There are no predictors, only group mean is fixed
                  + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
```

```
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PV1)                                ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PV1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.05813679   6% variance between schools
########################### Descriptive analyses and Null model for student level
variables##############################################################
# Student variable "ST004D01T"      (Student (Standardized) Gender)   1= Female ,   2= male
dim(removed1)
removed1$ST004D01T
summary(removed1$ST004D01T)
mean(removed1$ST004D01T)     # 1.492193
sd(removed1$ST004D01T)       # 0.4999607
skewness(removed1$ST004D01T) # 0.03122806
NULL.MODEL.gender <- lme4::lmer(ST004D01T ~ 1                   ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.gender)                             ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.gender),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
```

```
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)       # 0.04382903
#### ST097Q01TA
removed1$ST097Q01TA     # How often during <test language lessons>: Students don't listen to what the teacher says.
# Student variable ST097Q01TA (Frequency of not listening to what teacher says)  1=every lesson ,    4=never
summary(removed1$ST097Q01TA)  # OK min 1 , max 4
table(removed1$ST097Q01TA)    # OK 1,2,3,4
mean(removed1$ST097Q01TA)     # 1.678001
sd(removed1$ST097Q01TA)       # 0.7868498
skewness(removed1$ST097Q01TA) # 1.178935
##########ICC Check
NULL.MODEL.097.1 <- lme4::lmer(ST097Q01TA ~ 1                              ## There are no predictors, only group mean is fixed
                         + (1 | CNTSCHID),                           ## Denotes that the variable of interest is nested in schools
                         data = removed1,                            ## Specifies the dataset used in the analysis
                         REML=F,
                         control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                               optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.097.1)                                            ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.097.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)  #0.0475729
### ST097Q02TA
# How often during <test language lessons>: There is noise and disorder.
# Student variable ST097Q02TA (Frequency of noise and disorder)
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST097Q02TA)  # OK min 1 , max 4
table(removed1$ST097Q02TA)    # OK 1,2,3,4
```

```
mean(removed1$ST097Q02TA)     # 1.593685
sd(removed1$ST097Q02TA)       # 0.7304801
skewness(removed1$ST097Q02TA) # 1.262908
##########ICC Check
NULL.MODEL.097.2 <- lme4::lmer(ST097Q02TA ~ 1                                    ## There are no predictors, only group mean is fixed
                       + (1 | CNTSCHID),                              ## Denotes that the variable of interest is nested in schools
                       data = removed1,                              ## Specifies the dataset used in the analysis
                       REML=F,
                       control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                         optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.097.2)                                            ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.097.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05477586
#ST097Q03TA
# Student variable ST097Q03TA (Frequency of noise and disorder)
# How often during <test language lessons>: The teacher waits long for students to quiet down.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST097Q03TA)  # OK min 1 , max 4
table(removed1$ST097Q03TA)    # OK 1,2,3,4
mean(removed1$ST097Q03TA)     # 1.493754
sd(removed1$ST097Q03TA)       # 0.7514799
skewness(removed1$ST097Q03TA) # 1.634017
##########ICC Check
NULL.MODEL.097.3 <- lme4::lmer(ST097Q03TA ~ 1                                    ## There are no predictors, only group mean is fixed
                       + (1 | CNTSCHID),                              ## Denotes that the variable of interest is nested in schools
                       data = removed1,                              ## Specifies the dataset used in the analysis
                       REML=F,
```

```
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.097.3)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.097.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)  # 0.0466966
# Student variable ST097Q04TA (Frequency of noise and disorder)
# How often during <test language lessons>: Students cannot work well.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST097Q04TA)  # OK min 1 , max 4
table(removed1$ST097Q04TA)    # OK 1,2,3,4
mean(removed1$ST097Q04TA)     # 1.494795
sd(removed1$ST097Q04TA)       # 0.7144415
skewness(removed1$ST097Q04TA) # 1.52902
##########ICC Check
NULL.MODEL.097.4 <- lme4::lmer(ST097Q04TA ~ 1                       ## There are no predictors, only group mean is fixed
                        + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                        data = removed1,                   ## Specifies the dataset used in the analysis
                        REML=F,
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.097.4)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.097.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
```

```
print(ICC.between)     # 0.03084069
# Student variable ST097Q05TA (Frequency of noise and disorder)
# How often during <test language lessons>: Students don't start working for a long time after the lesson begins.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST097Q05TA)  # OK min 1 , max 4
table(removed1$ST097Q05TA)    # OK 1,2,3,4
mean(removed1$ST097Q05TA)     # 1.445958
sd(removed1$ST097Q05TA)       # 0.7115605
skewness(removed1$ST097Q05TA) # 1.729314
##########ICC Check
NULL.MODEL.097.5 <- lme4::lmer(ST097Q05TA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                     ## Denotes that the variable of interest is nested in schools
                    data = removed1,                     ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                     optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.097.5)                                    ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.097.5),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.04019553
# Student variable ST100Q01TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher shows an interest in every student's learning.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST100Q01TA)  # OK min 1 , max 4
table(removed1$ST100Q01TA)    # OK 1,2,3,4
mean(removed1$ST100Q01TA)     # 3.191187
sd(removed1$ST100Q01TA)       # 0.8872708
```

```
skewness(removed1$ST100Q01TA) # -0.8163337
##########ICC Check
NULL.MODEL.100.1 <- lme4::lmer(ST100Q01TA ~ 1                              ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.100.1)                                    ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.100.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)    # 0.04439306
dim(removed1)        # 11528    68
# Student variable ST100Q02TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher gives extra help when students need it
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST100Q02TA)  # OK min 1 , max 4
table(removed1$ST100Q02TA)    # OK 1,2,3,4
mean(removed1$ST100Q02TA)     # 3.31662
sd(removed1$ST100Q02TA)       # 0.8421086
skewness(removed1$ST100Q02TA) # -0.9950781
##########ICC Check
NULL.MODEL.100.2 <- lme4::lmer(ST100Q02TA ~ 1                              ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
```

```
summary(NULL.MODEL.100.2)                              ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.100.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between) # 0.02600348
dim(removed1)  # 11528    68
# Student variable ST100Q03TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher helps students with their learning.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST100Q03TA)  # OK min 1 , max 4
table(removed1$ST100Q03TA)    # OK 1,2,3,4
mean(removed1$ST100Q03TA)     # 3.453591
sd(removed1$ST100Q03TA)       # 0.7827836
skewness(removed1$ST100Q03TA) # -1.344124
##########ICC Check
NULL.MODEL.100.3 <- lme4::lmer(ST100Q03TA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                    data = removed1,                  ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                              optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.100.3)                              ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.100.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   #0.03455873
```

```
dim(removed1)
# Student variable ST100Q04TA (Frequency of the teacher showing an interest in every student's learning.)
# How often during <test language lessons>: The teacher continues teaching until the students understands.
# 1=every lesson , 2=Most lessons, 3=Some lessons  4=never or hardly ever
summary(removed1$ST100Q04TA)  # OK min 1 , max 4
table(removed1$ST100Q04TA)    # OK 1,2,3,4
mean(removed1$ST100Q04TA)     # 3.366586
sd(removed1$ST100Q04TA)       # 0.8297835
skewness(removed1$ST100Q04TA) # -1.141514
##########ICC Check
NULL.MODEL.100.4 <- lme4::lmer(ST100Q04TA ~ 1                       ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.100.4)                               ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.100.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   #0.02664582
dim(removed1)
removed1$ST211Q01HA
summary(removed1$ST211Q01HA)  # OK min 1 , max 4
table(removed1$ST211Q01HA)    # OK 1,2,3,4
mean(removed1$ST211Q01HA)     # 2.878557
sd(removed1$ST211Q01HA)       # 0.920826
skewness(removed1$ST211Q01HA) # -0.7655217
##########ICC Check
```

```
NULL.MODEL.211.1 <- lme4::lmer(ST211Q01HA ~ 1                          ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
                    data = removed1,                     ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.211.1)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.211.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.03212568
dim(removed1)
removed1$ST211Q02HA
summary(removed1$ST211Q02HA)  # OK min 1 , max 4
table(removed1$ST211Q02HA)    # OK 1,2,3,4
mean(removed1$ST211Q02HA)     # 2.816967
sd(removed1$ST211Q02HA)       # 0.9020667
skewness(removed1$ST211Q02HA) # -0.6626003
##########ICC Check
NULL.MODEL.211.2 <- lme4::lmer(ST211Q02HA ~ 1                          ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
                    data = removed1,                     ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.211.2)                                  ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.211.2),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.02888411
dim(removed1)
removed1$ST211Q03HA
summary(removed1$ST211Q03HA)  # OK min 1 , max 4
table(removed1$ST211Q03HA)    # OK 1,2,3,4
mean(removed1$ST211Q03HA)     # 2.891135
sd(removed1$ST211Q03HA)       # 0.9217862
skewness(removed1$ST211Q03HA) # -0.7519377
##########ICC Check
NULL.MODEL.211.3 <- lme4::lmer(ST211Q03HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.211.3)                         ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.211.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.03084168
dim(removed1)
removed1$ST213Q01HA
```

```
summary(removed1$ST213Q01HA)  # OK min 1 , max 4
table(removed1$ST213Q01HA)    # OK 1,2,3,4
mean(removed1$ST213Q01HA)     # 3.010236
sd(removed1$ST213Q01HA)       # 0.8145885
skewness(removed1$ST213Q01HA) # -0.8420409
##########ICC Check
NULL.MODEL.213.1 <- lme4::lmer(ST213Q01HA ~ 1                ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),              ## Denotes that the variable of interest is nested in schools
                    data = removed1,               ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.213.1)                          ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.213.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05813679
dim(removed1)
removed1$ST213Q02HA
summary(removed1$ST213Q02HA)  # OK min 1 , max 4
table(removed1$ST213Q02HA)    # OK 1,2,3,4
mean(removed1$ST213Q02HA)     # 2.931558
sd(removed1$ST213Q02HA)       # 0.8325073
skewness(removed1$ST213Q02HA) # -0.6427074
##########ICC Check
NULL.MODEL.213.2 <- lme4::lmer(ST213Q02HA ~ 1                ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),              ## Denotes that the variable of interest is nested in schools
```

```
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.213.2)                                  ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.213.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.0557366
dim(removed1)
removed1$ST213Q03HA
summary(removed1$ST213Q03HA)  # OK min 1 , max 4
table(removed1$ST213Q03HA)    # OK 1,2,3,4
mean(removed1$ST213Q03HA)     # 3.044067
sd(removed1$ST213Q03HA)       # 0.7646442
skewness(removed1$ST213Q03HA) # -0.9103795
##########ICC Check
NULL.MODEL.213.3 <- lme4::lmer(ST213Q03HA ~ 1                        ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.213.3)                                  ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.213.3),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.02824168
dim(removed1)
removed1$ST213Q04HA
summary(removed1$ST213Q04HA)  # OK min 1 , max 4
table(removed1$ST213Q04HA)    # OK 1,2,3,4
mean(removed1$ST213Q04HA)     # 3.024722
sd(removed1$ST213Q04HA)       # 0.8148633
skewness(removed1$ST213Q04HA) # -0.8196991
##########ICC Check
NULL.MODEL.213.4 <- lme4::lmer(ST213Q04HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                            ## Denotes that the variable of interest is nested in schools
                    data = removed1,                             ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.213.4)                                        ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.213.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.03173709
dim(removed1)
removed1$ST034Q01TA
```

```
summary(removed1$ST034Q01TA)  # OK min 1 , max 4
table(removed1$ST034Q01TA)    # OK 1,2,3,4
mean(removed1$ST034Q01TA)     # 1.973282
sd(removed1$ST034Q01TA)       # 0.8523194
skewness(removed1$ST034Q01TA) # 0.7376838
##########ICC Check
NULL.MODEL.34.1 <- lme4::lmer(ST034Q01TA ~ 1                ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),           ## Denotes that the variable of interest is nested in schools
                    data = removed1,            ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                              optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.34.1)                        ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.34.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.008415705
dim(removed1)
removed1$ST034Q04TA
summary(removed1$ST034Q04TA)  # OK min 1 , max 4
table(removed1$ST034Q04TA)    # OK 1,2,3,4
mean(removed1$ST034Q04TA)     # 2.030968
sd(removed1$ST034Q04TA)       # 0.796994
skewness(removed1$ST034Q04TA) # 0.6652473
##########ICC Check
NULL.MODEL.34.4 <- lme4::lmer(ST034Q04TA ~ 1                ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),           ## Denotes that the variable of interest is nested in schools
```

```
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                      optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.34.4)                                    ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.34.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.005836506
dim(removed1)
removed1$ST034Q06TA
summary(removed1$ST034Q06TA)  # OK min 1 , max 4
table(removed1$ST034Q06TA)    # OK 1,2,3,4
mean(removed1$ST034Q06TA)     # 1.965822
sd(removed1$ST034Q06TA)       # 0.853782
skewness(removed1$ST034Q06TA) # 0.7359338
##########ICC Check
NULL.MODEL.34.6 <- lme4::lmer(ST034Q06TA ~ 1                          ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                      optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.34.6)                                    ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.34.6),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.0100449
dim(removed1)
removed1$ST123Q02NA
summary(removed1$ST123Q02NA)  # OK min 1 , max 4
table(removed1$ST123Q02NA)    # OK 1,2,3,4
mean(removed1$ST123Q02NA)     # 3.25321
sd(removed1$ST123Q02NA)       # 0.865155
skewness(removed1$ST123Q02NA) # -1.231348
##########ICC Check
NULL.MODEL.123.2 <- lme4::lmer(ST123Q02NA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.123.2)                              ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.123.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.03228436
dim(removed1)
removed1$ST123Q03NA
```

```
summary(removed1$ST123Q03NA)  # OK min 1 , max 4
table(removed1$ST123Q03NA)    # OK 1,2,3,4
mean(removed1$ST123Q03NA)     # 3.208796
sd(removed1$ST123Q03NA)       # 0.8294499
skewness(removed1$ST123Q03NA) # -1.054571
##########ICC Check
NULL.MODEL.123.3 <- lme4::lmer(ST123Q03NA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.123.3)                    ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.123.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.01962145
dim(removed1)
removed1$ST123Q04NA
summary(removed1$ST123Q04NA)  # OK min 1 , max 4
table(removed1$ST123Q04NA)    # OK 1,2,3,4
mean(removed1$ST123Q04NA)     # 3.196131
sd(removed1$ST123Q04NA)       # 0.8425829
skewness(removed1$ST123Q04NA) # -1.047739
##########ICC Check
NULL.MODEL.123.4 <- lme4::lmer(ST123Q04NA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
```

```
                         data = removed1,                         ## Specifies the dataset used in the analysis
                         REML=F,
                         control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                        optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.123.4)                                     ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.123.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.02889342
dim(removed1)        # 11528    68
removed1$ST205Q01HA
summary(removed1$ST205Q01HA)  # OK min 1 , max 4
table(removed1$ST205Q01HA)    # OK 1,2,3,4
mean(removed1$ST205Q01HA)     #  2.549705
sd(removed1$ST205Q01HA)       #  0.8500644
skewness(removed1$ST205Q01HA) # -0.1596081
##########ICC Check
NULL.MODEL.205.1 <- lme4::lmer(ST205Q01HA ~ 1                          ## There are no predictors, only group mean is fixed
                         + (1 | CNTSCHID),                         ## Denotes that the variable of interest is nested in schools
                         data = removed1,                         ## Specifies the dataset used in the analysis
                         REML=F,
                         control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                        optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.205.1)                                     ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.205.1),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05432459
dim(removed1)
removed1$ST205Q02HA
summary(removed1$ST205Q02HA)  # OK min 1 , max 4
table(removed1$ST205Q02HA)    # OK 1,2,3,4
mean(removed1$ST205Q02HA)     #  2.550659
sd(removed1$ST205Q02HA)       #  0.856515
skewness(removed1$ST205Q02HA) #  -0.1583017
##########ICC Check
NULL.MODEL.205.2 <- lme4::lmer(ST205Q02HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                    ## Denotes that the variable of interest is nested in schools
                    data = removed1,                     ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.205.2)                              ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.205.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.06051184
dim(removed1)
removed1$ST205Q03HA
```

```
summary(removed1$ST205Q03HA)  # OK min 1 , max 4
table(removed1$ST205Q03HA)    # OK 1,2,3,4
mean(removed1$ST205Q03HA)     #  2.469466
sd(removed1$ST205Q03HA)       #  0.8732082
skewness(removed1$ST205Q03HA) #  -0.09452782
##########ICC Check
NULL.MODEL.205.3 <- lme4::lmer(ST205Q03HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.205.3)                              ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.205.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05732888
dim(removed1)
removed1$ST205Q04HA
summary(removed1$ST205Q04HA)  # OK min 1 , max 4
table(removed1$ST205Q04HA)    # OK 1,2,3,4
mean(removed1$ST205Q04HA)     # 2.545021
sd(removed1$ST205Q04HA)       # 0.9045078
skewness(removed1$ST205Q04HA) # -0.1405523
##########ICC Check
NULL.MODEL.205.4 <- lme4::lmer(ST205Q04HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
```

```
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution

summary(NULL.MODEL.205.4)                                    ## Need to extract variance components from the null model

Var.comp <- print(VarCorr(NULL.MODEL.205.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.04546036
dim(removed1)        # 11528    68
removed1$ST062Q01TA
summary(removed1$ST062Q01TA)  # OK min 1 , max 4
table(removed1$ST062Q01TA)    # OK 1,2,3,4
mean(removed1$ST062Q01TA)     # 1.68338
sd(removed1$ST062Q01TA)       # 0.8414902
skewness(removed1$ST062Q01TA) # 1.190399
##########ICC Check
NULL.MODEL.62.1 <- lme4::lmer(ST062Q01TA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.62.1)                                     ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.62.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
```

```
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.03066181
dim(removed1)
removed1$ST062Q02TA
summary(removed1$ST062Q02TA)  # OK min 1 , max 4
table(removed1$ST062Q02TA)    # OK 1,2,3,4
mean(removed1$ST062Q02TA)     # 1.740545
sd(removed1$ST062Q02TA)       # 0.8481973
skewness(removed1$ST062Q02TA) # 1.056665
##########ICC Check
NULL.MODEL.62.2 <- lme4::lmer(ST062Q02TA ~ 1                    ## There are no predictors, only group mean is fixed
                  + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                  data = removed1,                  ## Specifies the dataset used in the analysis
                  REML=F,
                  control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                 optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.62.2)                          ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.62.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.02452506
dim(removed1)
removed1$ST062Q03TA
summary(removed1$ST062Q03TA)  # OK min 1 , max 4
table(removed1$ST062Q03TA)    # OK 1,2,3,4
mean(removed1$ST062Q03TA)     # 1.704892
sd(removed1$ST062Q03TA)       # 0.841269
```

```
skewness(removed1$ST062Q03TA) # 1.125162
##########ICC Check
NULL.MODEL.62.3 <- lme4::lmer(ST062Q03TA ~ 1                    ## There are no predictors, only group mean is fixed
                 + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                 data = removed1,                  ## Specifies the dataset used in the analysis
                 REML=F,
                 control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                 optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.62.3)                               ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.62.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.04107475
dim(removed1)       # 11528    68
removed1$ST038Q04NA
summary(removed1$ST038Q04NA)  # OK min 1 , max 4
table(removed1$ST038Q04NA)    # OK 1,2,3,4
mean(removed1$ST038Q04NA)     # 1.445697
sd(removed1$ST038Q04NA)       # 0.8115591
skewness(removed1$ST038Q04NA) # 1.754029
##########ICC Check
NULL.MODEL.38.4 <- lme4::lmer(ST038Q04NA ~ 1                    ## There are no predictors, only group mean is fixed
                 + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                 data = removed1,                  ## Specifies the dataset used in the analysis
                 REML=F,
                 control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                 optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.38.4)                               ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.38.4),comp="Variance")
```

```
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.01520859
dim(removed1)
removed1$ST038Q05NA
summary(removed1$ST038Q05NA)  # OK min 1 , max 4
table(removed1$ST038Q05NA)    # OK 1,2,3,4
mean(removed1$ST038Q05NA)     # 1.337092
sd(removed1$ST038Q05NA)       # 0.7375266
skewness(removed1$ST038Q05NA) # 2.159185
##########ICC Check
NULL.MODEL.38.5 <- lme4::lmer(ST038Q05NA ~ 1                          ## There are no predictors, only group mean is fixed
                     + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                     data = removed1,                  ## Specifies the dataset used in the analysis
                     REML=F,
                     control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.38.5)                              ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.38.5),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.04854065
dim(removed1)


removed1$ST038Q06NA
```

```
summary(removed1$ST038Q06NA)  # OK min 1 , max 4
table(removed1$ST038Q06NA)    # OK 1,2,3,4
mean(removed1$ST038Q06NA)     # 1.365718
sd(removed1$ST038Q06NA)       # 0.7676176
skewness(removed1$ST038Q06NA) # 2.07665
##########ICC Check
NULL.MODEL.38.6 <- lme4::lmer(ST038Q06NA ~ 1                        ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                  optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.38.6)                              ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.38.6),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   #  0.04318189
dim(removed1)


removed1$ST038Q07NA
summary(removed1$ST038Q07NA)  # OK min 1 , max 4
table(removed1$ST038Q07NA)    # OK 1,2,3,4
mean(removed1$ST038Q07NA)     # 1.333796
sd(removed1$ST038Q07NA)       # 0.7488201
skewness(removed1$ST038Q07NA) # 2.212752
##########ICC Check
NULL.MODEL.38.7 <- lme4::lmer(ST038Q07NA ~ 1                        ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
```

```
                        data = removed1,                        ## Specifies the dataset used in the analysis
                        REML=F,
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.38.7)                                        ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.38.7),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05131053
dim(removed1)       # 11528    68


removed1$ST206Q01HA
summary(removed1$ST206Q01HA)  # OK min 1 , max 4
table(removed1$ST206Q01HA)    # OK 1,2,3,4
mean(removed1$ST206Q01HA)     # 2.832755
sd(removed1$ST206Q01HA)       # 0.8002316
skewness(removed1$ST206Q01HA) # -0.5231501
##########ICC Check
NULL.MODEL.206.1 <- lme4::lmer(ST206Q01HA ~ 1                   ## There are no predictors, only group mean is fixed
                        + (1 | CNTSCHID),                       ## Denotes that the variable of interest is nested in schools
                        data = removed1,                        ## Specifies the dataset used in the analysis
                        REML=F,
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.206.1)                                       ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.206.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
```

```
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05834301
dim(removed1)


removed1$ST206Q03HA
summary(removed1$ST206Q03HA)  # OK min 1 , max 4
table(removed1$ST206Q03HA)    # OK 1,2,3,4
mean(removed1$ST206Q03HA)     # 2.893737
sd(removed1$ST206Q03HA)       # 0.7664128
skewness(removed1$ST206Q03HA) # -0.523865
##########ICC Check
NULL.MODEL.206.3 <- lme4::lmer(ST206Q03HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                ## Denotes that the variable of interest is nested in schools
                    data = removed1,                 ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.206.3)                            ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.206.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.057394
dim(removed1)


removed1$ST206Q04HA
```

```
summary(removed1$ST206Q04HA)  # OK min 1 , max 4
table(removed1$ST206Q04HA)    # OK 1,2,3,4
mean(removed1$ST206Q04HA)     # 2.867713
sd(removed1$ST206Q04HA)       # 0.7922623
skewness(removed1$ST206Q04HA) # -0.5056807
##########ICC Check
NULL.MODEL.206.4 <- lme4::lmer(ST206Q04HA ~ 1                                ## There are no predictors, only group mean is fixed
                        + (1 | CNTSCHID),                          ## Denotes that the variable of interest is nested in schools
                        data = removed1,                           ## Specifies the dataset used in the analysis
                        REML=F,
                        control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                      optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.206.4)                                    ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.206.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.06022867
dim(removed1)       # 11528    68


removed1$ESCS
summary(removed1$ESCS)  # OK min - 4.5977  , max 3.9991
table(removed1$ESCS)    # OK 1,2,3,4
mean(removed1$ESCS)     # -0.2338557
sd(removed1$ESCS)       # 0.8427418
skewness(removed1$ESCS) # -0.1957476
##########ICC Check
NULL.MODEL.ESCS <- lme4::lmer(ESCS ~ 1                                ## There are no predictors, only group mean is fixed
                       + (1 | CNTSCHID),                          ## Denotes that the variable of interest is nested in schools
```

```
                    data = removed1,                    ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                       optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.ESCS)                                      ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.ESCS),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.234107
dim(removed1)        # 11528    68
PV1.5.Logic <- colnames(removed1) %in% c("PV1MATH", "PV2MATH", "PV3MATH","PV4MATH","PV5MATH")
PVMMATH <- apply(removed1[, PV1.5.Logic], 1, FUN=function(x)mean(x, na.rm=T))
PV2.5.Logic <- colnames(removed1) %in% c("PV1SCIE", "PV2SCIE", "PV3SCIE","PV4SCIE","PV5SCIE")
PVMSCIE <- apply(removed1[, PV2.5.Logic], 1, FUN=function(x)mean(x, na.rm=T))
PV3.5.Logic <- colnames(removed1) %in% c("PV1READ", "PV2READ", "PV3READ","PV4READ","PV5READ")
PVMREAD <- apply(removed1[, PV3.5.Logic], 1, FUN=function(x)mean(x, na.rm=T))
removed1 <- cbind.data.frame(removed1, PVMMATH, PVMSCIE, PVMREAD)
dim(removed1) # 11528    71
removed1$PVMMATH
summary(removed1$PVMMATH)  # min 188.0,  max 741.1
table(removed1$PVMMATH)    #
mean(removed1$PVMMATH)     # 455.9054
sd(removed1$PVMMATH)       # 85.02674
skewness(removed1$PVMMATH) # 0.1620681
##########ICC Check
NULL.MODEL.PVMMATH <- lme4::lmer(PVMMATH ~ 1                        ## There are no predictors, only group mean is fixed
                     + (1 | CNTSCHID),                   ## Denotes that the variable of interest is nested in schools
                     data = removed1,                    ## Specifies the dataset used in the analysis
                     REML=F,
```

```
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PVMMATH)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PVMMATH),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   #0.3779237
dim(removed1)
removed1$PVMSCIE
summary(removed1$PVMSCIE)  # min 197.2  ,  max 711.6
table(removed1$PVMSCIE)    #
mean(removed1$PVMSCIE)     # 431.0671
sd(removed1$PVMSCIE)       # 82.12551
skewness(removed1$PVMSCIE) # 0.4819615
##########ICC Check
NULL.MODEL.PVMSCIE <- lme4::lmer(PVMSCIE ~ 1                 ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),             ## Denotes that the variable of interest is nested in schools
                    data = removed1,              ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PVMSCIE)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PVMSCIE),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]

total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
```

```
print(ICC.between)   # 0.4129872
dim(removed1)
removed1$PVMREAD
summary(removed1$PVMREAD)  # min 183.3 ,  max 716.0
table(removed1$PVMREAD)    #
mean(removed1$PVMREAD)     # 421.6316
sd(removed1$PVMREAD)       # 85.6875
skewness(removed1$PVMREAD) # 0.3717105
##########ICC Check
NULL.MODEL.PVMREAD <- lme4::lmer(PVMREAD ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                 ## Denotes that the variable of interest is nested in schools
                    data = removed1,                 ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                              optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.PVMREAD)                       ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.PVMREAD),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.3990201
dim(removed1)        # 11528    71
#207.1
removed1$ST207Q01HA
summary(removed1$ST207Q01HA)  # OK min 1 , max 4
table(removed1$ST207Q01HA)    # OK 1,2,3,4
mean(removed1$ST207Q01HA)     # 2.746096
sd(removed1$ST207Q01HA)       # 0.970182
skewness(removed1$ST207Q01HA) # -0.5357426
##########ICC Check
```

```
NULL.MODEL.207.1 <- lme4::lmer(ST207Q01HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.207.1)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.207.1),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.06016467
dim(removed1)      # 11528    71
#207.2
removed1$ST207Q02HA
summary(removed1$ST207Q02HA)  # OK min 1 , max 4
table(removed1$ST207Q02HA)    # OK 1,2,3,4
mean(removed1$ST207Q02HA)     # 2.904382
sd(removed1$ST207Q02HA)       # 0.8689248
skewness(removed1$ST207Q02HA) # -0.7488731
##########ICC Check
NULL.MODEL.207.2 <- lme4::lmer(ST207Q02HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                    optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.207.2)                                  ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.207.2),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
```

```
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05265747
dim(removed1)      # 11528    71
#207.3
removed1$ST207Q03HA
summary(removed1$ST207Q03HA)  # OK min 1 , max 4
table(removed1$ST207Q03HA)    # OK 1,2,3,4
mean(removed1$ST207Q03HA)     # 2.92644
sd(removed1$ST207Q03HA)       # 0.9307422
skewness(removed1$ST207Q03HA) # -0.7173636
##########ICC Check
NULL.MODEL.207.3 <- lme4::lmer(ST207Q03HA ~ 1                ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),            ## Denotes that the variable of interest is nested in schools
                    data = removed1,             ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                              optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.207.3)                        ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.207.3),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.07024156
dim(removed1)       # 11528    71

#207.4
removed1$ST207Q04HA
```

```
summary(removed1$ST207Q04HA)  # OK min 1 , max 4
table(removed1$ST207Q04HA)    # OK 1,2,3,4
mean(removed1$ST207Q04HA)     # 2.841777
sd(removed1$ST207Q04HA)       # 0.9026243
skewness(removed1$ST207Q04HA) # -0.6153265
##########ICC Check
NULL.MODEL.207.4 <- lme4::lmer(ST207Q04HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
                    data = removed1,                   ## Specifies the dataset used in the analysis
                    REML=F,
                    control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                                   optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.207.4)                              ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.207.4),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.04742501
dim(removed1)       # 11528    71

#207.5
removed1$ST207Q05HA
summary(removed1$ST207Q05HA)  # OK min 1 , max 4
table(removed1$ST207Q05HA)    # OK 1,2,3,4
mean(removed1$ST207Q05HA)     # 2.929216
sd(removed1$ST207Q05HA)       # 0.9011256
skewness(removed1$ST207Q05HA) # -0.7774587
##########ICC Check
NULL.MODEL.207.5 <- lme4::lmer(ST207Q05HA ~ 1                    ## There are no predictors, only group mean is fixed
                    + (1 | CNTSCHID),                  ## Denotes that the variable of interest is nested in schools
```

```
                 data = removed1,                    ## Specifies the dataset used in the analysis
                 REML=F,
                 control = lme4::lmerControl(optimizer = "optimx", calc.derivs = FALSE,
                              optCtrl = list(method = "nlminb")))  # these are special controls for finding the optimal solution
summary(NULL.MODEL.207.5)                                   ## Need to extract variance components from the null model
Var.comp <- print(VarCorr(NULL.MODEL.207.5),comp="Variance")
Var.comp <- as.data.frame(Var.comp)
Between.School.var <- Var.comp$vcov[1]
Within.school.var <- Var.comp$vcov[2]
total.var <- Between.School.var + Within.school.var
ICC.between <- Between.School.var/total.var
print(ICC.between)   # 0.05322191
dim(removed1)       # 11528    71

###### School-level descriptive analyses######
# Location
dim(removed1)
removed1$SC001Q01TA
summary(removed1$SC001Q01TA)  # min -2.4195 ,  max 1.5805
table(removed1$SC001Q01TA)    #
mean(removed1$SC001Q01TA)     # -0.00000000000000001826487
sd(removed1$SC001Q01TA)       #  1.288426
skewness(removed1$PVMREAD)    #  0.3717105
### School type
dim(removed1)
removed1$priv.only
summary(removed1$priv.only)  # min 1 ,  max 2
table(removed1$priv.only)    # 10967  ,  561
mean(removed1$priv.only)     # 1.048664
sd(removed1$priv.only)       # 0.2151742
skewness(removed1$priv.only)    #  4.19471
### School type
```

```
dim(removed1)
removed1$part.only
summary(removed1$part.only)  # min 1 ,  max 2
table(removed1$part.only)    # 11057,  471
mean(removed1$part.only)     # 1.040857
sd(removed1$part.only)       # 0.1979675
skewness(removed1$part.only) # 4.638166


### AVG.PISA.ESCS
dim(removed1)
removed1$AVG.PISA.ESCS
summary(removed1$AVG.PISA.ESCS)  # min -1.11407 ,  max 0.97367
table(removed1$AVG.PISA.ESCS)    #
mean(removed1$AVG.PISA.ESCS)     # -0.000000000000000001048267
sd(removed1$AVG.PISA.ESCS)       # 0.458417
skewness(removed1$AVG.PISA.ESCS) # 0.2020207
### SCHSIZE
dim(removed1)
removed1$SCHSIZE
summary(removed1$SCHSIZE)  # min -1.5012   ,  max 3.8481
table(removed1$SCHSIZE)    #
mean(removed1$SCHSIZE)     #  0.000000000000000008539089
sd(removed1$SCHSIZE)       # 1
skewness(removed1$SCHSIZE) # 1.264311
#### CREACTIV
dim(removed1)
removed1$CREACTIV
summary(removed1$CREACTIV)  # min 0.000   ,  max 3.000
table(removed1$CREACTIV)    # OK 0,1,2,3
mean(removed1$CREACTIV)     #  2.159004
sd(removed1$CREACTIV)       # 0.9699907
```

```
skewness(removed1$CREACTIV) # -0.7547581
#### EDUSHORT
dim(removed1)
removed1$EDUSHORT
summary(removed1$EDUSHORT)  # min -1.3244   ,  max 3.0563
table(removed1$EDUSHORT)    #
mean(removed1$EDUSHORT)     #  -0.00000000000000003311488
sd(removed1$EDUSHORT)       # 1.119071
skewness(removed1$EDUSHORT) # 0.3507112
#### STAFFSHORT
dim(removed1)
removed1$STAFFSHORT
summary(removed1$STAFFSHORT)  # min -1.0985   ,  max 4.5222
table(removed1$STAFFSHORT)    #
mean(removed1$STAFFSHORT)     # -0.00000000000000002904088
sd(removed1$STAFFSHORT)       # 0.9709163
skewness(removed1$STAFFSHORT) # 0.8538587
#### W_FSTUWT
dim(removed1)         # 11528    71
removed1$W_FSTUWT
removed1$W_FSTUWT <- removed1$W_FSTUWT - mean(removed1$W_FSTUWT)
mean(removed1$W_FSTUWT)     # 0.00000000000000003655811
summary(removed1$W_FSTUWT)  # min -8.718    ,  max 31.761
table(removed1$W_FSTUWT)    #
mean(removed1$W_FSTUWT)     #  0.00000000000000003655811
sd(removed1$W_FSTUWT)       # 7.520449
skewness(removed1$W_FSTUWT) # 1.423897
dim(removed1) # 11528    71
###################### Table 1 analysis##############################################
apply(removed1, 2, function(x)min(x))     # average of variable (min)
apply(removed1, 2, function(x)max(x))     # average of variable (min)
apply(removed1, 2, function(x)mean(x))     # average of variable (mean)
```

```
apply(removed1, 2, function(x)sd(x))     # average of variable (sd)
apply(removed1, 2, function(x)skewness(x))     # average of variable (skewness)
#### Descriptive Statistics for Within-School (Student-Level) Variables ####
# Calculate Design Effect to examine necessity for MLM
length(unique(removed1$CNTSCHID))  #426 schools   with 11528 students
dim(removed1) # 11528    71
# Design Effect = 1 + (m - 1) * ICC
11528/426   # Average number of students per school 27.06103
def<- 1+(27.06103-1)*c(0.044,0.048,0.055,0.047,0.031,0.040,0.044,0.026,0.035,
                    0.027,0.032,0.029,0.031,0.058,0.056,0.028,0.032,0.008,
                    0.006,0.010,0.032,0.020,0.029,0.054,0.061,0.057,0.045,
                    0.031,0.025,0.041,0.015,0.049,0.043,0.051,0.058,0.057,
                    0.060,0.060,0.053,0.070,0.047,0.053,0.234,0.378,0.413,0.399)
print(def)
length(which(def>2))    # 28 variables with DEFF of more than 2.
#################Note, some negative values and anomolous results so removed schools with fewer than 10 students, then re-ran model
# Removing students with fewer than 10 students
removed1<-removed1[with(removed1, CNTSCHID %in% names(which(table(CNTSCHID)>=10))), ]
dim(removed1)      # 11317    71
sort(table(removed1$CNTSCHID))
length(unique(removed1$CNTSCHID))  #399 schools   with 11317 students
dim(removed1)       #11317    71
length(unique(removed1$CNTSCHID))  #399 schools
model <- '
        ST038w =~ ST038Q04NA + ST038Q05NA + ST038Q06NA + ST038Q07NA
        ST207w =~ ST207Q01HA + ST207Q02HA + ST207Q03HA + ST207Q04HA + ST207Q05HA
        ST097w =~ ST097Q01TA + ST097Q02TA + ST097Q03TA + ST097Q04TA + ST097Q05TA
        ST213w =~ ST213Q01HA + ST213Q02HA + ST213Q03HA + ST213Q04HA
        ST100w =~ ST100Q01TA + ST100Q02TA + ST100Q03TA + ST100Q04TA
        ST211w =~ ST211Q01HA + ST211Q02HA + ST211Q03HA
        ST206w =~ ST206Q01HA + ST206Q03HA + ST206Q04HA
        ST205w =~ ST205Q01HA + ST205Q02HA + ST205Q03HA + ST205Q04HA
```

```
        ST034w =~ ST034Q01TA + ST034Q04TA + ST034Q06TA
        ST123w =~ ST123Q02NA + ST123Q03NA + ST123Q04NA
        ST62w  =~ ST062Q01TA + ST062Q02TA + ST062Q03TA
         '
fit <- cfa(model = model,
          data = removed1,
          std.lv=T,
          estimator="MLR")
summary(fit, fit.measures=TRUE)
estim <- parameterestimates(fit, standardized=TRUE)
options(max.print=1000000)
print(estim)
########################Inter factor correlation matrix for the measurement model
lambda<-c(0.703,0.841,0.609)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)
### 038. Student experience of bullying   AVE
lambda<-c(0.785,0.921,0.896,0.924)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                  # AVE 0.7802645
### 207. Anti-bullying attitude
lambda<-c(0.814,0.875,0.846,0.861,0.868)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                  # AVE 0.7277364
### 097. Disruptive Student Academic Behavior   AVE
lambda<-c(0.693,0.735,0.718,0.597,0.719)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                  # AVE 0.4818736
# 213.Teacher enthusiasm
lambda<-c(0.840,0.834,0.826,0.846)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                  #AVE 0.699787
```

```
#100. Teacher's support and teaching practices
lambda<-c(0.665,0.782,0.819,0.740)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                #AVE 0.5680275
#211Teacher behavior and student learning
lambda<-c(0.875,0.853,0.857)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                #AVE 0.742561
###206. Student Cooperation AVE
lambda<-c(0.844,0.935,0.850)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                #
# AVE= 0.769687
###205. Student Competition
lambda<-c(0.822,0.901,0.884,0.728)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                 #
# AVE= 0.6997313
###034. Lack of Student Sense of Belonging
lambda<-c(0.738,0.803,0.806)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                #
# AVE= 0.6130297
###123. Parental involvement in school activities
lambda<-c(0.896,0.924,0.868)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)                #
# AVE= 0.8033387
###062. Student Truancy and Lateness
lambda<-c(0.703,0.841,0.609)
AVE<-sum(lambda^2)/(sum(lambda^2)+sum(1-(lambda^2)))
print(AVE)
```

```
# AVE= 0.5241237
### Ten school climate factors were included in the study.
# Student Truancy and Lateness were removed from the study
# due to Related inter-factor variance (r2) > AVE
#038. Student experience of bullying
#207. Anti-bullying attitude
#097. Disruptive Student Academic Behavior
#213.Teacher enthusiasm
#100. Teacher's support and teaching practices
#211Teacher behavior and student learning
#206. Student Cooperation
#205. Student Competition
#034. Lack of Student Sense of Belonging
#123. Parental involvement in school activities
dim(removed1)   #11317   71
#### Checking all school climate factors without ESCS and gender variables
model <- '
    ST038w =~ ST038Q04NA + ST038Q05NA + ST038Q06NA + ST038Q07NA
    ST207w =~ ST207Q01HA + ST207Q02HA + ST207Q03HA + ST207Q04HA + ST207Q05HA
    ST097w =~ ST097Q01TA + ST097Q02TA + ST097Q03TA + ST097Q04TA + ST097Q05TA
    ST213w =~ ST213Q01HA + ST213Q02HA + ST213Q03HA + ST213Q04HA
    ST100w =~ ST100Q01TA + ST100Q02TA + ST100Q03TA + ST100Q04TA
    ST211w =~ ST211Q01HA + ST211Q02HA + ST211Q03HA
    ST206w =~ ST206Q01HA + ST206Q03HA + ST206Q04HA
    ST205w =~ ST205Q01HA + ST205Q02HA + ST205Q03HA + ST205Q04HA
    ST034w =~ ST034Q01TA + ST034Q04TA + ST034Q06TA
    ST123w =~ ST123Q02NA + ST123Q03NA + ST123Q04NA

    PVMMATH ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w
    PVMSCIE ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w
    PVMREAD ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w
   '
```

```
fit <- sem(model = model,
        data = removed1,
        std.lv=T,
        estimator="MLR")
summary(fit, fit.measures=TRUE, rsquare=T)
estim <- parameterestimates(fit, standardized=TRUE)
options(max.print=1000000)
print(estim)
# f-squared value = .R^2/ (1-R^2)
#PVMMATH        0.141
#PVMSCIE        0.167
#PVMREAD        0.196
.141/ (1-.141)  # 0.1641444   meduim
.167/ (1-.167)  # 0.2004802   medium
.196/ (1-.196)  # 0.2437811   meduim
#Effect on PVMMATH 0.1641444
#Effect on PVMSCIE 0.2004802
#Effect on PVMREAD 0.2437811
################Checking all school climate factors together with gender and ESCS variables
model <- '

      ST038w =~ ST038Q04NA + ST038Q05NA + ST038Q06NA + ST038Q07NA
      ST207w =~ ST207Q01HA + ST207Q02HA + ST207Q03HA + ST207Q04HA + ST207Q05HA
      ST097w =~ ST097Q01TA + ST097Q02TA + ST097Q03TA + ST097Q04TA + ST097Q05TA
      ST213w =~ ST213Q01HA + ST213Q02HA + ST213Q03HA + ST213Q04HA
      ST100w =~ ST100Q01TA + ST100Q02TA + ST100Q03TA + ST100Q04TA
      ST211w =~ ST211Q01HA + ST211Q02HA + ST211Q03HA
      ST206w =~ ST206Q01HA + ST206Q03HA + ST206Q04HA
      ST205w =~ ST205Q01HA + ST205Q02HA + ST205Q03HA + ST205Q04HA
      ST034w =~ ST034Q01TA + ST034Q04TA + ST034Q06TA
      ST123w =~ ST123Q02NA + ST123Q03NA + ST123Q04NA
```

```
	PVMMATH ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w + ST004D01T + ESCS
	PVMSCIE ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w +  ST004D01T + ESCS
	PVMREAD ~ ST038w + ST207w + ST097w + ST213w + ST100w + ST211w + ST206w + ST205w + ST034w + ST123w +  ST004D01T + ESCS
	'
fit <- sem(model = model,
	data = removed1,
	std.lv=T,
	estimator="MLR")
summary(fit, fit.measures=TRUE, rsquare=T)
estim <- parameterestimates(fit, standardized=TRUE)
options(max.print=1000000)
print(estim)
# f-squared value = .R^2/ (1-R^2)
#PVMMATH 	0.184
#PVMSCIE 	0.204
#PVMREAD 	0.223
.184/ (1-.184)  # 0.2254902   meduim
.204/ (1-.204)  # 0.2562814   medium
.223/ (1-.223)  # 0.2870013   meduim
#Effect on PVMMATH 0.2254902
#Effect on PVMSCIE 0.2562814
#Effect on PVMREAD 0.2870013
```