UNIVERSITY OF GOTHENBURG

FACULTY OF EDUCATION
DEPARTMENT OF EDUCATION AND SPECIAL EDUCATION

THE MEDIATING ROLE OF STUDENT'S ACADEMIC SELF-BELIEFS IN STUDYING EDUCATIONAL EQUITY

A comparison between Sweden and Ukraine in TIMSS 2011

Oleksandra Mittal

Master's thesis:	30 credits
Programme/course:	L2EUR (IMER) PDA184
Level:	Second cycle (Advanced)
Term/year:	Autumn 2019
Supervisor:	Kajsa Yang Hansen
Examiner:	Ilse Hakvoort

Abstract

Master's thesis:	30 credits
Programme/Course:	L2EUR (IMER) PDA184
Level:	Second cycle (Advanced)
Term/year:	Autumn 2019
Supervisor:	Kajsa Yang Hansen
Examiner:	Ilse Hakvoort
Keywords:	Academic self-beliefs, Science self-concept, Interest and Utility, Socioeconomic status mediators, Structural Equation Modelling with multiple mediators, Educational Equity, Sweden, Ukraine, TIMSS 2011

Aim: The aim of this study is to enrich the understanding of student-level mechanisms involved in shaping (in)equity in educational outcomes among eight-graders in Sweden and Ukraine by examining the role of academic self-beliefs in science, represented by motivational constructs of self-concept, interest and utility value, as mediators of the association between student socioeconomic status (SES) and achievement in science.

Theory: Expectancy-value theory (EVT) of achievement motivation (Eccles (Parsons) et al., 1983) laid out the fundament for the author's empirical model. Their perspective was used to identify within TIMSS 2011 two sets of hierarchical self-beliefs constructing academic identity. To further establish them as the mediators between SES and academic achievement and thus build a conceptual framework, the reference to Eccles' EVT model, Bandura's sociocognitive perspective (2012) and Krapohl & Plomin's (2015) finding on genetical mediation along with a literature study were used.

Method: Structural Equation Modeling (SEM) is applied in this study to TIMSS 2011 data for Sweden and Ukraine using Mplus version 7.4 software. Only single-level model with complex option is fit into data for reporting results, as the author's focus is on student-level mediating constructs. Intraclass correlation coefficients are represented to support this choice. It was initially planned that five latent indices would be used within the specified model, however, one latent mediator representing student's interest in science was excluded due to multicollinearity issues. Measurement invariance was performed to check for parameter estimates' comparability across Sweden and Ukraine.

Results: Only academic self-concept or self-confidence in science was found to be a significant non-cognitive mediating factor in the interplay between student socioeconomic status and science achievement in both Sweden and Ukraine. In Sweden, the total indirect effect of two academic self-beliefs makes up for ¼ of SES – science achievement relationship, whereas academic self-concept alone mediates one third of this relationship. In Ukraine, half of SES – science achievement association is mediated by the effect of academic self-concept. Students' value of science (utility) had a non-significant indirect effect, although it added up to self-concept indirect effect in Ukraine and reduced the total indirect effect in Sweden by .01 points. Interest had to be excluded from the structural model due to its high correlation with both measures discussed above. The finding confirms that in order to get a fair image of educational equity in the country it is important to go from the individual level up to family, classroom, school and society, and no other way around, as they are additional layers built up around the individual's non-cognitive and cognitive characteristics.

Acknowledgements

This thesis has been a long journey for me, but no traveller can make it without the favourable circumstances and support from people who come in our lives. Therefore, I would like to express my gratitude to everyone who made this thesis possible.

At first, I would like to thank my supervisor Prof. Kajsa Yang-Hansen for accepting all my ideas with openness and encouragement, for providing support with the thesis at any time of night and day required and replying on my numerous e-mails. The field of empirical research in education was pretty new to me when I entered the IMER programme and I feel proud to have mastered it at least a little and to have such a brilliant teacher.

I also want to thank every tutor in IMER who opened a whole world of new perspectives in educational research and were always supportive and understanding. My extended thanks go to our programme coordinator Dr. Ernst Thoutenhoofd, who sincerely cared for the progress of each student and was there to offer his immense knowledge of education field whenever I needed an advice or had a doubt.

At a time in the past, education field was rather chosen for me than by me, but now I want to thank my dear parents for placing me in the right path, as now I cannot think of the righter one for me. My enormous gratitude goes to my parents-in-law, or simply just mummy and papa, without whose support the idea of getting International Master's in Sweden wouldn't have been possible and have remained just an idea.

And finally, I would like to thank two most important people in my life, my husband, Mr. Shubham Mittal and my daughter Iryna. Because of you, Shubham, I'm sitting today and writing the foreword in my Master thesis, with your incredible support and belief in me I dared to come to Sweden to do my master's and continued till the end. Your own perseverance, confidence and inexhaustible strength and energy inspire me always and I promise not to stop in advancing further and use my potential to the fullest. Iryna, I don't have much hopes that you will read this thesis completely when you grow up, but I want you to know that you have been my motivation all the way through my studies and writing this work. Since the last five years you have changed me in the ways no one could, you have strengthened and polished my character, you are and will always be my biggest support. I strive that I always be someone you will be proud of when you grow up, and this thesis is the first step towards that ideal.

Table of Contents

List of Tables and Figures

List of Abbreviations

TIMSS – Trends in International Mathematics and Science Study, conducted by IEA since 1995

IEA – International Association for the Evaluation of Educational Achievement

PIRLS – Progress in International Reading Literacy Study, conducted by IEA

PISA – Programme for International Student Assessment, conducted by OECD

OECD – Organization for Economic Cooperation and Development

ILSA – International Large-Scale Assessment

UNDP – United Nations Development Programme

SES – socioeconomic status

EVT – Expectancy-Value Theory

Introduction

According to the report of Swedish National Agency for Education for the year 2009, two years prior to TIMSS [1] 2011 – the most scrutinized survey with regards to studying educational equity – was conducted, the Swedish school had put the responsibility for learning on students, by using investigative forms of teaching and different forms of "own work", since the late 1990s (Skolverket, 2009). Thus, different individual factors, cognitive and non-cognitive, started playing more significant role for Swedish students' achievement. On the other hand, deteriorating trends have been observed in educational equity in Sweden since 1980s and until 2014: despite the decentralization and free school choice policies, aimed at reducing the relationship between socioeconomic status (SES) and achievement, this association only increased (Yang Hansen & Gustafsson, 2018). Educational equity is traditionally measured through examining the level of correlation between student socioeconomic status and academic achievement, with the low correlation signifying the equity (White, 1982; Sirin, 2005). In the quest of explaining how the SES-achievement relationship works, there has been an emerging trend in developing explanatory models for this association – mediating and moderating (Gustafsson, Nilsen & Yang Hansen, 2018; Liu et al, 2015; Rjosk et al., 2014; Guo et al., 2018). Beside a number of studies done on school factors, students' characteristics are more and more often identified as playing an important role in explaining the correlation between social background and academic achievement (e.g. Johnson, McGue, & Iacono, 2007; Steinmayr, Dinger, & Spinath, 2010). However, very few studies have been conducted in Sweden, examining this mediating relationship (see Mood et al., 2012), and revealing how individual prerequisites mediate the effects of socioeconomic status on school achievement helps to understand one of the mechanisms steering the formation of (in)equity in education.

Ukraine, which achievement in science and mathematics in TIMSS 2011 survey administered to eighth grade was at the same level with Sweden, has been underrepresented in the field of international educational research to a great extent, with only rare occasional qualitative studies appearing with regards to its education system, and none to the researcher's knowledge existing quantitative research, although data is being collected. The only recent study mentioning Ukraine was the one by Gustafsson, Nilsen and Yang Hansen (2018) who analyzed the total of 50 countries-participants of TIMSS 2011 for school characteristics that have potential to reduce the association between SES and academic achievement, and Ukraine was in the list of countries which educational system may be compensatory in the relation to student SES. Nevertheless, the question of educational equity has not been investigated for the country with the population of 42 million people – the biggest in the Eastern Europe. Therefore, it is high time to bring it in the spotlight of educational research by investigating how student-level characteristics mediate SES – academic achievement relationship on the example of two countries which students perform at par with each other, but which educational systems are so different.

[1] Trends in International Mathematics and Science Study, conducted by IEA.

Within this thesis the term of "equity" and not "equality in education" is used with the explanation of both the concepts given further. "Equality of educational opportunity" was the fundamental reason for the establishment of public education in Europe and Northern America in the nineteenth century, and was about the resources put into education, equal rights, universal education and sameness in treating all (Volckmar, 2018). The concept of "equity" was introduced by Coleman (1968), and it replaced the race for equal educational opportunities. Educational equity or "equity for equal achievement" (Espinoza, 2007) has been an important agenda for the countries around the globe for many decades now and is one of the six levers of education policy in Sweden (OECD, 2013; 2017). Equity can be understood as fairness in the education provision, in which comparatively same educational outcomes are achieved by the means of unequal distribution of resources in order to address individual circumstances and prerequisite skills and needs of students. Thereby, the differences in input do not matter as long as students are able to fulfill their potential and reach more or less same level of academic performance. Equity and equality – two sides of the coin – are sometimes mistakenly interchanged, however, for the purpose of understanding the given research it is important to emphasize that the author will use the term "equity".

This study aims to enrich the understanding of the growing inequity in educational outcomes in Sweden by examining the role of science self-concept, interest and utility value – two sets of self-beliefs shaping academic identity in science domain (Eccles, 2009), – as mediators of the association between student socioeconomic status and achievement in science. To give a solid explanation of this mediating relationship, the model will be replicated for Ukraine that will be used as a means of comparing and explaining results in Sweden. This will as well bring Ukraine at the front of international educational research concerned with educational equity and mechanisms related to it for the first time.

To understand the mediating mechanism studied here, it is important for the reader to know that mediating variables explain the relationship between the independent and dependent variables, with complete mediation resulting in the initial predictor variable no longer affecting the outcome. Mediators must fulfill two requirements: be casually related to the predictor and the outcome (Baron & Kenny, 1986). There has been a number of studies concluding that student psychological variables – cognitive and non-cognitive – are the best predictors of academic achievement (Marzano, 2000; ten studies in Hattie, 2009; Stankov et al., 2012; Stankov & Lee, 2017). To explain the relationship between students' socioeconomic status and psychological factors as well as to reason the choice of three non-cognitive constructs, the Eccles & Wigfield expectancy-value model of achievement motivation is used as a theoretical framework (2002), besides the reference is given to Bandura's sociocognitive perspective (2012) and Krapohl & Plomin's (2015) finding on genetically influenced factors mediating SES – academic achievement association. Therefore, two requirements for making a mediating model are achieved, where the literature study on the theoretical perspectives and previous findings in this area is built around outlining and discussing all the relationships implied by the mediating mechanism. All of it builds a comprehensive conceptual framework.

TIMSS 2011 assessment administered to eighth-grade students in Sweden and Ukraine is going to be analyzed for this purpose. Capturing how the hierarchy of academic self-beliefs mediate SES – academic achievement relationship in TIMSS survey is important, as TIMSS is a curriculum-based assessment programme, and its conclusions about educational equity affect policies directly rather than, for example, PISA[2] which measures student abilities as applied in real-life situations without restricting to the formal curriculum (OECD, 2016). Structural Equation Modeling is chosen as the analytical method which has a long-established reputation of a preferred technique in analyzing causal relationships using international large-scale assessments (ILSAs).

The study is important for policymakers and educational psychologists – to see what role student psychological effects play in shaping (in)equity, and for teachers – to encourage them to apply differential techniques for each student to enable everyone to reach the highest potential and, thus, to improve educational equity.

The thesis starts with providing the rationale for choosing Sweden and Ukraine for the analysis purpose by including a short overview of their education systems and comparison within the TIMSS framework, as well as outlining the matter of educational equity. It is then followed by the Literature Study introducing the mediation mechanism, expectancy-value theory along with other perspectives and findings, all of which shape a conceptual framework supporting the researcher's stand and establishing a reliable fundament for the work. Academic self-concept is highlighted as the main mediating factor. Definitions of key terms are provided. Research questions and hypotheses conclude the first part of the thesis followed by the Methodology, which outlines all the steps taken in the second part of the thesis. Limitations and ideas for further research are discussed in the end of this work after Results and Conclusions sections.

Background

Educational Equity in Sweden and Ukraine

In Sweden, "all children and young persons shall have equal access to education, irrespective of gender, geographic residence, or financial circumstances" (Mullis, 2012). Its teachers and school leaders work on adapting education to the needs of students thus inclusivity is the priority for all educational institutions. Main characteristics of the Swedish education system are free school choice introduced since the beginning of 1990s, and high decentralization with municipalities deciding on the state subsidy allocation to schools and taking all responsibilities for meeting the educational goals set by Skolverket – Swedish National Agency for Education. Sweden has been known to invest lavishly in educational sector being among the OECD leaders in education spending (OECD, 2003), although the expenditure somewhat decreased since the 1990s (Helgøy & Homme, 2006). According to Beach et al. (2013), the segregation between schools has increased since 2006 with some schools comprising of 100% of

[2] Programme for International Student Assessment, conducted by OECD.

students with immigrant backgrounds, which may lead to ethnic inequality and social exclusion of youth.

With regard to educational equity in Sweden, it is an important research topic here with researchers not detecting any change in SES – academic performance relationship since 1990s, until a publication by Gustafsson and Yang Hansen (2018) has given an overview of the impact family background has had on student achievement since 1988 until 2014. According to their findings, there has been a slight increase of the correlation between parental education and achievement by 0.4 units. Additionally, student characteristics like age and ethnicity were found to moderate the relation between the SES and achievement. An attempt to explain this issue was made by analyzing the relationship between parental education and achievement on three levels: student, school and municipality; besides, the moderating effect of school characteristics was studied in the relationship between SES and school achievement based on TIMSS 2011 data (Gustafsson, Nilsen & Yang Hansen, 2018).

By the Ukrainian Education law, the chief of any educational institution has to guarantee that no discrimination or privileged treatment takes place on the grounds of race, color, political, religious and other beliefs, gender, age, disability, ethnic origin, family and socioeconomic status, place of residence, language (retrieved and translated from the official government's website https://zakon.rada.gov.ua/laws/show/2145-19). Inclusive education is guaranteed by the state to those with special educational needs, although in practice it is not always possible due to the lack of school resources. Ukraine represents a more ethnically homogeneous population with a highly centralized public education system. Its Ministry of Education and Science, Youth and Sport is responsible for almost all the matters related to education, major and minor: educational policies, standards and curricula development, specifying school subjects to be studied, certifying course books and materials and providing all public school students with course books, conducting the final state examination and independent external assessment as the obligatory condition for enrollment in higher education institutions, in-service teacher training (Mullis, 2012). In contrast to Sweden, it is not a wealthy state with all its education infrastructure established during the times of its membership in the Soviet Union, and a complex post-Soviet transition is still going on since 2000s. Striving to tighten relations with EU, it joined the Bologna process in 2005 and introduced an array of educational reforms (Filiatreau, S., 2011). At the year of TIMSS 2011 assessment new sets of national educational standards for mathematics and science were developed and approved by the Cabinet of Ministers of Ukraine (Mullis et al., 2012). Unfortunately, it is impossible as of now to see the effect of those reforms as Ukraine has not participated in TIMSS since 2011 due to the focus being shifted to political agendas. However, it did participate in PISA 2018, and this work could be a good fundament to continue empirical research on Ukraine.

The choice to take Sweden and Ukraine for the comparative analysis was motivated by a few significant factors. If to compare them within TIMSS framework, both Sweden and Ukraine' performance in mathematics and science was at par with each other. Sweden's average achievement was 484 in mathematics and 509 in science, while Ukraine scored 480 in mathematics and 501 in science (see *Figure 1* for science achievement). Apart from that, Sweden experienced drastic fall in science performance since the first TIMSS assessment in 1995, during the years students' achievement steadily fell from 553 to 509 in the year 2011. Ukraine participated only two times in TIMSS 2007 and 2011, and its performance went up from 485 to 501 (Martin et al., 2012). After comparing school factors in both countries (see *Table 1*), Ukraine gives way to Sweden as it has more problems with educational resources at schools, and the schools on average are composed of 59% of disadvantaged students, while Swedish schools' are represented mostly by affluent students (74%). Ukrainian students also have less home resources, however, the level of parental education is almost equal for both states.

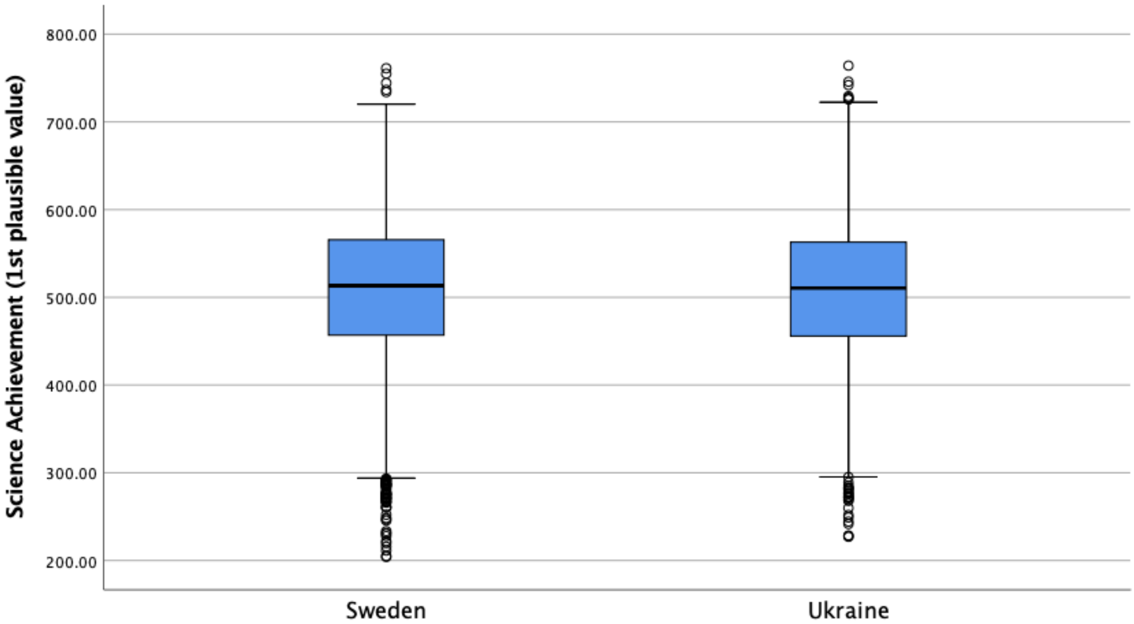


Figure 1. Distribution of Science Achievement in Sweden and Ukraine, TIMSS 2011.

With regard to teachers, in both the countries teachers have to be highly-educated. Five-year teacher training at a higher education institution is a must for Ukrainian teachers, with four years formally titled as bachelor's degree usually considered to be incomplete higher education. In Sweden, teacher education takes 3,5-5 years and students have freedom to choose the combination of subjects to be studied which Ukrainian institutions do not provide. In same manner, Ukrainian teachers are very restricted with regards to choosing teaching methods and topics to be covered in lessons, while teaching materials are decided by the state and only particular schools can deviate from that in individual subjects. According to TIMSS 2011 survey among science and mathematics teachers of eighth graders, 93% of such teachers in Ukraine are women which can be explained by the comparatively low remuneration and a cultural factor that teaching is attributed to as a women's profession, while in Sweden this distribution of science and mathematics teachers at the eighth-grade level is almost equal.

As of year 2011, children in Sweden used to begin school in the autumn of the year when they turned seven years old, and in Ukraine compulsory schooling begins at age 6. As a rule, Ukrainian parents play an active part in their child's school life and usually help their children with homework which is given since the first grade. Sometimes schooling is a joint venture where parents track their child's academic success, communicate with teachers, while participation and even organization of school activities by parents is a common thing. This leads to the suggestion that parents may play a significant role in shaping their children's academic self-concept. In Sweden, the culture of parental involvement is less extensive, students are freer in their school life and get no grades until the sixth grade. Emphasis on homework has been increased starting from lower secondary school since Swedish students' poor performance in PISA 2009 and 2012. This resulted in parents more and more often choosing the services of private companies that support students' homework (Gu & Kristoffersson, 2015). Apart from that, parents play a leading role in selecting a school for their children which can be based on their child's abilities, socioeconomic or ethnical background, leading to ability stratification among schools.

To conclude, different mechanisms might be involved in shaping educational (in)equity and student achievement in both the countries, but individual-level mechanism is the first step to see what role a student plays from educational psychology point of view in mediating parents' socioeconomic status impact on educational outcomes: is he or she more influenced by the external factors or does personality come first?

Table 1. Comparison of Ukraine and Sweden based on UNDP report (2014), and data from TIMSS 2011

Factors	Sweden	Ukraine
Population total (2011)	9 481 000	45 477 689
Government type	Parliamentary constitutional monarchy	Semi-presidential republic
Net migration rate (2010/2015)	4.2 migrant(s)/1,000 people	-0.2 migrant(s)/1,000 people
Gross National Income (GNI) per capita (PPP) (2013)	$43,201	$8,215
Human Development Index - HDI (UNDP, 2013)	0.898 (12th)	0.734 (83rd)
Gini coefficient (2003-2012) – income inequality	25.0	25.6
Inequality-adjusted education index (2013)	0.800	0.747
School composition by students' economic background:		
- More affluent[a]	74	13
- Balanced composition	21	29
- More disadvantaged[b]	5	59
Percent of students in urban and rural areas:		
- with more than 100,000 residents	22	31
- 15,001 to 100,000 residents	42	18
- 15,000 or fewer	36	52
Percentage of female teachers	53	92.5
School emphasis on academic success, %		
- very high	5	0
- high	62	31
Instruction affected by science resource shortage, %		
- not affected	40	1
- somewhat affected	60	80
- affected a lot	0	19

Safe and orderly school, %		
- safe and orderly	29	63
- somewhat	67	37
- not safe and orderly	4	0
Parents' highest education, %		
- University or higher	48	44
- Post-secondary, not University	28	34
- Upper Secondary	18	15
Home educational resources, %		
- many resources	27	12
- some resources	71	79
Components of home educational resources, %:		
- more than 100 books at home	42	25
- own room and Internet connection at home	94	48
- at least one parent with a university degree or higher	47	39

Literature Study

This section represents a comprehensive conceptual framework, - a summation of all the findings and perspectives the author uses to support the relationships among the factors claimed in the designed and proposed by her individual-level mediation model (*figure 2*), along with the main theoretical point of departure for further setting of the research questions and hypotheses. Prior to that, the readers are reminded once again on what is mediation in order to understand the logic behind structuring literature review the way it is.

Mediation Mechanism

The mediation model is used in this work in order to examine one of the unobservable mechanisms by which student socioeconomic status affects science achievement. It can also be referred to as a black box model (Weed, 1998), where the black box includes all the unobservable variables through which the predictor affects the outcome or stimulus-organism-response (S-O-R) model which was first introduced in behavioural sciences in 1928 by Woodworth outlining that mental processes can translate the stimulus effect on response (MacKinnon, 2008). Baron & Kenny's landmark article (1986) on mediators and moderators distinguished the properties of mediating variables in social science research, and it is their definition and schematic representation of mediation that is most widely used by the educational researchers:

"Mediation is the generative mechanism through which the focal independent variable is able to influence the dependent variable of interest" (Baron & Kenny, 1986).

According to Wu & Zumbo (2007), mediation is a unidirectional trivariate hypothesis and a theory used to refine and understand a causal relationship. The questions of "How" and "Why" X predicts Y are answered by the mediation model. The *figure 2* given below represents a hypothesized by the researcher mediation mechanism which includes eighth grader's science self-beliefs that answer the questions of how and why student SES affects science achievement. Additional control variables added to the model are discussed in the Methodology section. Further discussion provided.

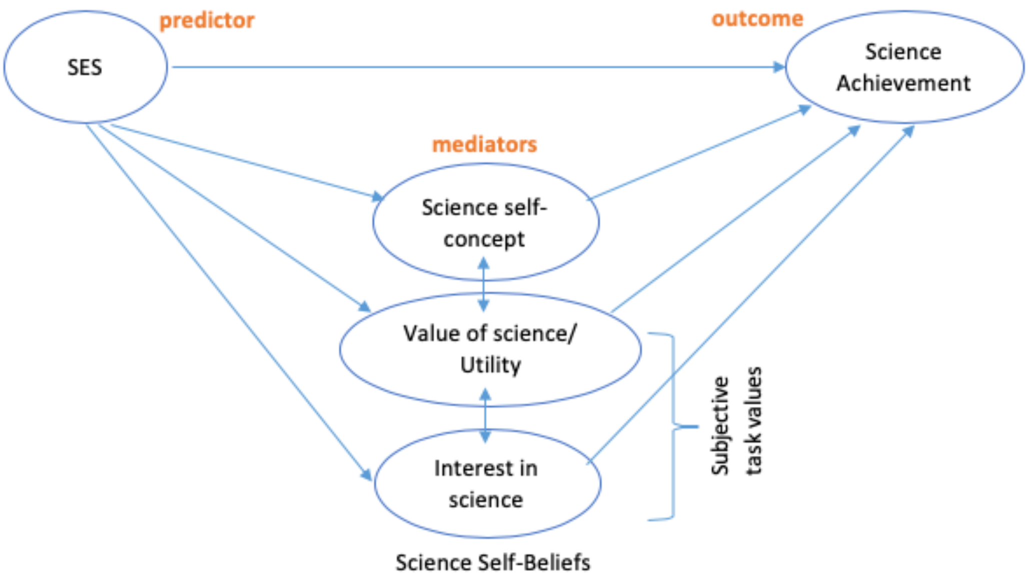


Figure 2. The hypothesized model of the individual-level mediating mechanism of educational equity.

Academic Self-Concept and Expectancy-Value Theory

Academic self-concept stands for one's current belief or perception about the ability to perform in a particular academic domain based on interpreted experiences gained in the environment and through interaction with others (Gasa et al., 2018). It is a facilitator of success in educational settings (Marsh & Craven, 2006), and a key psychological construct. The reader should remember to differentiate it from self-efficacy concept that means one's perceived confidence in their future performance evaluated by a student on the basis of his or her self-concept. Academic self-concept is sometimes regarded as domain-related self-confidence (Sheldrake, 2016), which is also true for the TIMSS 2011 measure of self-concept titled as *Students Confident in Science*.

Self-concept is a multidimensional structure which hierarchical model was first proposed by Shavelson et al. (1976) who divided it into global self-concept and academic domain-related self-concept, with its construct validity thoroughly scrutinized (Wylie, 1989). Since then there has been a consistent interest in academic self-concept that has a high predictive power related to interest, academic achievement, occupational choices and long-term educational attainment (Craven & Marsh, 2008; Parker et al., 2014; Parker et al., 2018). Initially academic self-concept comprised of both the affect (interest) and competence components which represented unified factors until Eccles and her colleagues (1983) proposed their expectancy-value model with self-beliefs and task values representing separate but positively interrelated components of academic motivation (Arens et al., 2011). It was further proved by Marsh et al. (1999) that competence and affect components are not only theoretically but empirically distinguishable within each domain of academic self-concept.

Before proceeding with theoretical framework, it is important to mention one more empirical finding providing the rationale for using academic self-concept as the mediator. Krapohl & Plomin (2015) looked into genetical mediation between parents' socioeconomic status and student educational outcomes and found that two third of SES – achievement association was explained by variables other than intelligence. Thereby, the conclusion was reached that other genetically influenced factors like personality traits, self-efficacy or academic self-concept (Kriegbaum & Spinath, 2016) play a role in differences in student achievement and its association with socioeconomic status. This finding supports the choice of researcher to use academic self-concept as a mediating or intermediate variable.

Underpinning the study is the expectancy-value model of achievement motivation (Eccles (Parsons) et al., 1983; Eccles and Wigfield, 2002; Eccles, 2009). Expectancy-value theory (EVT) was the first to introduce the importance of academic self-beliefs, expectancies for success and subjective task values (interest and utility value). Their perspective has been used to identify within TIMSS 2011 two sets of hierarchical self-beliefs constructing academic identity in science and serving as a source of motivational contribution to science achievement:

- *academic self-concept* – the confidence in one's current ability to succeed and being able to engage in scientific practices – is represented by "Students Confident in Science" scale built on students' own interpretation of their abilities and level of anxiety, as well as the effects of contextual factors translated from peer- and subject-wise comparisons, and teacher encouragement (see *Appendix 2*). Consequently, science self-concept informs task values.

- *subjective task values* – the values related to the identity formation which a person attaches to various achievement-related options, which in turn influence their life-defining choices (Eccles, 2009) – represented in the analyzed ILSA by two components of *intrinsic value*/*interest* and *utility or extrinsic value. Intrinsic value* in TIMSS 2011 is measured with "Students Like Learning Science" scale and *utility value* is encompassed within the scale of "Students Value Science".

The theoretical model suggested for empirical research is mainly grounded on the supposition of EVT (Eccles (Parsons) et al., 1983; Wigfield & Eccles, 2000) that children' s socioeconomic status influences their academic achievement via expectations for success or academic self-beliefs and task values. Both of the concepts are also called as motivational prerequisites or factors. It is interesting to note that the association between the influence of parents, parents' level of education and home environment in particular, and the self-concept of adolescents was already pointed out in 1943 by Maslow's theory of motivation (1943). His proposed hierarchy of needs explained why one might have positive or negative self-concept, where lower needs are required to be obtained from the family and higher-level needs – from the school.

To support EVT model, a reference to Bandura's socio-cognitive perspective is also important, according to which students' self-beliefs and academic aspirations are shaped by parents' familial belief systems which are influenced by their socioeconomic status (2012; Bandura et al., 2001). In turn, children's confidence about being able to succeed in a particular domain has a motivational power to influence achievement. The value which a child attaches to a certain academic activity can be determined by ability self-beliefs and vice versa (Eccles, 2009). Thus, going back to science achievement, students' academic self-concept in science is interrelated with their intrinsic and extrinsic values of science, playing a big part in shaping a self-defined academic identity and motivation of a student. Consequently, these academic self-beliefs influence student achievement and are influenced by parents' socioeconomic status, which makes them an important SES mediating mechanism.

Socioeconomic Status and its Mediators

As per the big body of educational research, student family socioeconomic status (SES) remains one of the most influential factors in predicting academic achievement (White, 1982; Sirin, 2005). Hattie in his meta-analysis of 499 quantitative studies discovered that this relationship has the biggest effect size (d=.57), meaning that SES explained 57% of variance in academic achievement (2009). Consequently, the high level of association between SES and achievement signals about the educational inequity, which is studied nowadays within the framework of international large-scale student surveys, such as TIMSS, PISA or PIRLS[3] (OECD, 2014a). On the global scale, extensive educational reforms introduced across the countries have not minimized the positive relationship between socioeconomic status and educational outcomes, leading to a conclusion that educational equity has not improved (Marks, 2013, p.172). However, the linear relationship between SES and academic achievement is more complex than once defined by Coleman (1968) in the biggest of its kind report on Equality for Educational Opportunity in US, who was the first to empirically prove that family background was actually the main determinant of child's success at learning.

In order to understand the mechanisms underlying the association between socioeconomic status and educational outcomes, a number of studies have emerged since the last decade exploring mediating and moderating factors which can strengthen, weaken or explain this relationship. For example, Liu et al. investigated the mediating effects of school processes influencing the relationship between school SES and mathematic literacy (2015). School climate, instructional quantity and quality are the most common factors explored as mediating school and classroom SES effect on achievement (Rjosk et al., 2014). Gustafsson, Nilsen, & Yang Hansen examined the moderating power of these predictors within schools across 50 countries participating in TIMSS 2011 (2018). Few more studies in the beginning of 2000s examined the mediating power of underlying family processes in the relationship between SES and school performance (Chao & Willms, 2002; Guo & Harris, 2000).

A small number of academics in general have investigated the student characteristics, cognitive and non-cognitive, as the mediators of family's SES impact on educational outcomes. For example, the

[3] Progress in International Reading Literacy Study.

mediating role of personality and motivational constructs was studied in relation to academic and mathematical achievement of adolescents, which was found to be significant (Steinmayr et al. 2010; 2012; Kriegbaum & Spinath, 2016). Marsh and colleagues (2005) using a German PISA dataset discovered that psychological and institutional factors, such as intelligence, enjoyment of reading, decoding skills, and school type, partially mediated socioeconomic status and reading competence association. Hsin and Xie, in their longitudinal study of cognitive and non-cognitive factors of kindergarten age children mediating the parental SES effect on achievement, found that non-cognitive factors mediation effects increase over time (2017). Parental involvement and expectation were found as significant mediators of SES – reading achievement relationship among 4-6th graders in China (Guo et al., 2018). In one earlier study of Bandura and his colleagues (2001), academic self-efficacy (another type of self-belief) entirely mediated family socioeconomic status' effects on children's career aspirations.

As for Sweden, a lot of student-level factors as mediators remain unstudied. The only research known to the author is the one by Mood et al. (2012) who examines cognitive versus socio-behavioral skills as mediators of father's education influence on educational outcomes of their 18-year-old sons using Swedish registry data with approximately 180,000 fathers and sons. Only weak mediation by socio-behavioral traits like social maturity, emotional capacity and leadership skills was found but no explanation of this given by the researchers. Ukraine, as already mentioned before, has been an outcast in terms of educational research on SES, its mediators, and equity in general.

Apart from that, science domain usually remains in shadow when the question of educational equity is investigated, and mathematics is given preference instead. Nevertheless, science holds as important as mathematics for the future generations, and its more multifaceted content encompassing four subjects can give a more comprehensive image of educational equity within school systems. As a conclusion, there is a need for a study on the mediation capacity of individual non-cognitive factors, such as academic self-concept, intrinsic and extrinsic value, in relation to SES – science achievement association in Sweden and Ukraine. In the case of significant mediation effect and using the definition of equity presented in this study, it can be argued that in an equitable educational system better individual traits lead to better educational outcomes in science (Kriegbaum & Spinath, 2016). Then, it is possible to conclude that the goal of educational equity is achieved despite of positive association between SES and science achievement.

Academic Self-Beliefs and Student Achievement

The findings for the effect of self-beliefs, including self-concept and other non-cognitive constructs, are somewhat contradictive. For example, the meta-review of 55 reports (mostly U.S. samples) published between 1978 and 2001 concluded that the average relationship between self-beliefs and subsequent achievement was not very strong with β = .08 effect size, the effect size was larger for the non-U.S. samples (β = .11) which meets Cohen's threshold of r = .10 for a small effect size (Valentine, DuBois and Cooper, 2004). In Hattie's summary of over 800 meta-analyses, 10 studies out of 31 reporting

Cohen's effect size greater than 1.00 reveal student characteristics to have stronger impact on academic achievement than school or teacher factors with average *Cohen's d* = 1.48 (2009).

The predictive power of academic self-concept was studied in the meta-analysis of cross-sectional studies (Hansford & Hattie, 1982) and was found to be highly associated with achievement ($r = .42$) in contract with measures of global self-concept ($r = .18$). Correlations between task values and educational outcomes are usually about $r = .30$. (Kriegbaum, Jansen, & Spinath, 2015; Steinmayr & Spinath, 2009). According to Wigfield & Eccles research (2002), a strong predictive power of academic self-concept or belief about one's ability held even after controlling for previous achievement. Same results were observed with Marsh et al. (2005) study of academic self-concept influences among seventh-graders which was empirically concluded to predict both grades and standardized test scores.

The number of recent studies claim that confidence which is highly correlated with self-concept is the best non-cognitive predictor of academic achievement (Stankov et al., 2012; Stankov, Morony, & Lee, 2014; Sheldrake, 2016). And in the latest study measuring the relationship between SES and self-beliefs, self-beliefs at the individual level of analysis were found to predict mathematics achievement better than the SES indices for both TIMSS and PISA data (Lee & Stankov, 2017), and at the country-level, the measure of self-concept (but not self-efficacy and anxiety) and the index of home possessions as an SES measure had equally high predictive power for student achievement. These findings further confirm the mediating role of academic self-beliefs between SES and academic achievement, and highlight the importance of measuring students' non-cognitive factors as mediators in order to understand whether high association between SES and achievement can be actually interpreted as educational inequity or whether the underlying mechanisms like student characteristics partially explain the high association, reducing the direct relationship between SES and achievement, making an education system look equitable.

Academic Self-Beliefs and Socioeconomic Status

Empirical research of association between socioeconomic status and self-beliefs provide different information. Thus, early studies of Olowu (1986), and Trusty, Peck, & Mathews (1994), reported that students of high SES had significantly higher self-concept than those of low SES. Opposing these findings is a new study of Gasa et al. (2018) presenting no significant relationship between the self-concept and any of the three components of socioeconomic status. However, a limitation to this study could be a very small sample size of 31 learners only.

Further studies on the association between the SES and self-concept are: Chohan and Khan (2010), Vyas and Choudhary (2017) who confirmed that low socioeconomic status of parents may affect their children in a bad way resulting in low self-concept as needs of children are not satisfied. Kaur, Rana & Kaur (2009) also supported this view who suggested that the development of self-concept can depend on family dynamics as well as family SES.

Definitions of Key Terms

To summarize the concepts discussed in detail in the previous Background section of the thesis, short definitions for each are given:

Academic self-beliefs – non-cognitive factors; the hierarchy of one's beliefs in his or her academic abilities and motivational expressions constructing academic identity; all used in this thesis are related to the science domain:

- **science self-concept and self-confidence** – two concepts used interchangeably which represent academic self-concept – one's belief or confidence in current academic abilities;
- **interest** – intrinsic motivation, part of subjective task values;
- **utility or utility value** – extrinsic motivation, part of subjective task values.

TIMSS (Trends in International Mathematics and Science Study) – international large-scale assessment (ILSA), founded and conducted each four years by the International Association for the Evaluation of Educational Achievement (IEA) since 1995 among fourth- and eighth-graders across the countries. Detailed discussion on this survey is provided in Methodology section.

Socioeconomic status (SES) – examined here as a student-level social and economic status, family background represented by TIMSS 2011 measure of *Home Educational Resources* (see *Appendix 1*). For the ease of understanding, the author uses SES abbreviation in her empirical model but cautions that this scale represents socioeconomic status only partially, and awareness of this should be maintained when interpreting the results.

Educational equity – fairness in the education provision which accounts for student's individual circumstances and makes sure that each individual has an opportunity to fulfill its highest potential, whereas more or less even distribution of educational outcomes is achieved for all.

Mediator - "a variable that occurs in a causal pathway from an independent variable to a dependent variable. It causes variation in the dependent variable and itself is caused to vary by the independent variable" (Last, 1988). Thus, mediator transmits the effect of an independent variable on a dependent variable and the power of this mediation is estimated as mediating effect or indirect effect.

Research Questions and Hypotheses

The study is aimed at examining the relationship between student socioeconomic status and science achievement ***mediated by*** two sets of academic self-beliefs in science represented by students' evaluation of their present abilities (science self-concept or self-confidence), and their task values (interest and utility value of science). This relationship will be compared between two countries: Sweden and Ukraine. The research questions of the proposal are:

- Do students’ science self-beliefs mediate the relationship between the student socioeconomic status and science achievement in Sweden and Ukraine, and to what extent?
- If yes, is this mediation significant to partially compensate for inequalities at the students’ background level, resulting in equitable educational outcomes in science for all in both countries?

In order to answer these research questions, and based on the previous research that has been done, the following hypotheses have been formulated which will be tested for Sweden and Ukraine:

H1: Students’ SES is casually related to science self-concept, interest and utility, in turn, science-self-concept, interest and utility factors are casually related to science achievement.

H1.$_0$: There is no significant relationship among SES, science self-beliefs and science achievement constructs.

H2: Students’ science self-confidence significantly mediates the association between their SES and scientific competence, resulting in partial compensation for inequity at students’ background level.

H2.$_0$: The mediating or indirect effect of self-confidence is not significant to partially compensate for inequitable educational outcomes as a consequence of students’ SES.

H3: Students’ interest and value of science significantly mediate the association between their SES and science achievement, resulting in partial compensation for inequity at students’ background level.

H3$_0$: The indirect effect of science interest and utility are not significant to partially compensate for inequitable educational outcomes as a consequence of students’ socioeconomic status.

H4: The mediating effect of science self-concept will hold when simultaneously considering the mediating effects of interest and utility, and vice versa.

H4$_0$: The indirect effect of science self-concept will NOT hold when simultaneously considering the indirect effect of interest and utility, and vice versa.

H5: The mediating effects of science self-belief factors on SES – science achievement relationship will remain significant controlling for students’ gender, ethnicity and parental support variables.

H5$_0$: After controlling for students’ gender, ethnicity and parental support variables, the mediating effects of one or both science self-belief constructs will become NON-significant.

Methodology

This master thesis applied structural equation modeling technique to data from TIMSS 2011. Samples and variables, as well as analytical method and process will be presented in detail in this section.

To answer the research questions, single-level Structural Equation Modeling (SEM) with complex option was used. SEM is a statistical analysis technique that is effectively employed in studying causal mechanisms within the framework of international large-scale assessments (ILSAs) of students’ competencies, such as TIMSS, PISA and PIRLS (Muijs, 2012). Single-level complex SEM model was chosen since the research focus here is on student's personal traits and the mechanism among these traits affecting academic outcomes. The constructs involved refer to individual phenomena, and intraclass correlation was performed to support that.

SEM comprises of measurement models (i.e., confirmatory factor analysis – CFA) and structural models. A CFA model examines the relationship between a set of indicators and the underlying theoretical construct, the so-called latent variable. In such a model, the factor loadings capture the degree of relatedness of each indicator and the latent variable. The higher the factor loading is the better the indicator is related to the latent construct. Normally, a factor loading should be over .30 so that the indicator can be regarded as an appropriate measure to the latent variable (Brown, 2014). CFA models build the foundation of a structural model, where interrelationship among the latent constructs measured by the CFA models are estimated. CFA models are estimated and evaluated first and then put together according to the theoretical model to examine the mechanism among the constructs. In this thesis, a total of four latent constructs were measured by CFA and used later in the structural model – two mediators measuring science self-confidence and utility (SciConf and SciValue), one control variable measuring parental support (SUPPORT) and science achievement factor (TotScie). In order to proceed with comparing parameter estimates between Sweden and Ukraine, the measurement invariance was tested on three degrees, namely, scalar invariance, metric invariance and configural invariance. The good model fit indices for metric level invariance are the minimum requirement for the estimates to be comparable across selected groups. To conduct all the described steps of analysis, Mplus version 7.4 was utilized (Muthén & Muthén, 1998–2012), while SPSS software was used for preparing the data, descriptive statistics and cross-checking mediating effects.

Data Source and Sampling Strategy

TIMSS 2011 data from Sweden and Ukraine are used as empirical foundation for examining the research questions in this thesis. TIMSS 2011 was the fifth cycle of the large-scale comparative study of fourth- and eighth-grade student knowledge in the curriculum areas of mathematics and science, administered every four years by the International Association for the Evaluation of Educational Achievement (IEA) since 1995 (Martin et al., 2012). In the year 2011, 45 countries and 14 benchmarking participants with nationally representative samples of over 300,000 students in total administered eighth grade assessment within the TIMSS framework.

In TIMSS a two-stage stratified cluster sample design with a systematic random sampling approach is applied. At the first stage, the schools in the target population are arranged into groups or strata according to a number of demographic variables, such as geographic region, school type, language of instruction, socio-economic indicators etc., which can take two forms: explicit and implicit stratification. Then the school samples are drawn from each type of strata ensuring that estimates are equally reliable for each geographic region and each group of schools with different demographic characteristics is proportionally represented. This is *probability-proportional-to-size school sampling*. At the second stage, taking into account the minimum class size specification, one or more complete classes are selected from each participating school with equal probability of selection regardless of the school size using *WinW3S* sampling software. This is *inversely proportional sampling*, where there is higher chance for classes from small schools to be selected. These stages result in interdependency of individuals in one clustered sample, which cannot be treated as independent observations.

In each participating country school stratification works differently. As an example, in Sweden explicit stratification variables included percentage of immigrants in the area and newly opened schools, while in Ukraine schools were stratified explicitly only as per urban and rural areas. Implicit stratification information for Sweden included levels of student performance, and for Ukraine – distribution as per administrative centers (Foy et al., 2013).

Being essential to a two-stage cluster sample design, a unique student sampling weight is generated at three levels – school, class (within school), and student (within class), and consists of two components: the inverse of the probability of selection at that level, together with an adjustment for nonparticipation (Joncas & Foy, 2012). Because of unequal probabilities of selecting a school, a class and a student, sampling weights are used to avoid bias in the parameter estimates (Laukaityte & Wiberg, 2017), however, the same leads to the variability of the generated weights across the sampled schools. For example, the weights can be larger in a smaller school and vice versa, causing occasional disproportion between the student population size and student sampling weights. This can be considered as a limitation to the reliability of parameter estimates for the population.

This type of hierarchical sampling with students clustered, for example, as per the schools' geographic location or their socioeconomic factors, results in too high intra-class correlation (ICC) within groups which violates standard statistical tests' assumption of the independency of observations (Hox, Moerbeek, & van de Schoot, 2010). And if violated, too small standard error estimates will be computed which leads to the rejection of null hypothesis and false significant results in the population – type I error (Muijs, 2010). Therefore, in this thesis complex model option was applied with a multistage sample to adjust for standard error estimates, besides, intraclass correlation was computed to support the researcher's choice of one-level over the two-level analysis.

Study Sample and Preparation of Data

Our sample comprises of 5,573 Swedish eight-graders from 153 schools (age mean = 14.8), and 3,378 Ukrainian eight-graders from 148 schools (age mean = 14.2) which represented the populations of eighth-grade-students in Sweden and Ukraine as of year 2011. TIMSS uses UNESCO's International Standard Classification of Education (ISCED) to identify the levels of schooling, thus, eight years after the first year of ISCED Level 1 would be the target grade for eighth grade TIMSS which is applicable to most of the countries (Mullis et al., 2009). The data consist of students' responses to 217 science items and to the student background questionnaire.

Student files for Sweden and Ukraine were merged by utilizing the International database analyzer (IEA, 2012). Data for both countries from 2011 included only identical items. Additionally, prior to running models in Mplus, the data was cleaned in SPSS software with only important variables left and renamed for the simplicity of analysis, and a number of dummy variables created. The number of classes in our samples was 266 in Sweden, and 162 in Ukraine with the average of 21 students per class (*see Table 2*; for more factors comparing Ukraine and Sweden including data taken from 2014 UNDP report please see *Table 1*).

Table 2. Main facts about the participants in Sweden and Ukraine.

Participants Included in the Analysis	Sweden	Ukraine
N schools	153	148
N classes	266	162
N students	5573	3378
Average students per class	21	21
% of girls	48	50
Students' age in years (mean)	14.8	14.2

Assessment Framework

In the pursuit of a comprehensive student assessment within the areas of science and mathematics, TIMSS' developed 28 assessment blocks (14 with science items and 14 with mathematics items) for the eighth grade which are estimated to contain about 10½ hours of testing time. Being far too extensive for the one student to handle, these blocks are distributed across 14 student achievement booklets with only one booklet administered to each student-participant. In the eighth grade a student is given 90 minutes to complete the booklet with additional 30 minutes for a student questionnaire (Mullis et al., 2009). This type of assessment design is called multiple-matrix sampling where the greater part of data is missing. Item response theory (IRT) is used then to aggregate results for all the booklets and to provide results for the entire assessment. Consequently, multiple imputation method is applied in TIMSS to generate reliable achievement estimates in the population: five plausible values for each student are computed with the variability accounting for the imputation error (Foy, 2012).

The author's focus is on factors related with science performance, however, the scope of this study allows only for a short overview of TIMSS science assessment framework. It was developed in collaboration with participating countries, and includes four major content domains – biology, chemistry, physics and earth science, – and three cognitive domains of applying, reasoning and knowing which cover the thinking processes the students activate when engaging with different science content (Mullis et al., 2009). In both Sweden and Ukraine science is taught as four separate subjects, thus the latent construct for the overall student achievement in science was created using imputation technique, which will be explained later in this study.

Apart from measuring student content knowledge and cognitive skills, the contextual background questionnaires are an essential part of TIMSS providing an insight into the factors affecting students' learning (Mullis et al., 2009). In TIMSS 2011, the questionnaires are administered to students, teachers and principals. The curriculum questionnaire, in addition, collects information on educational policies and contexts across the countries.

Reliability and Validity of TIMSS 2011

There are three key concepts inherent to the measurement stage of every good quantitative research study – reliability, validity and generalizability (Muijs, 2004). *Reliability* tells us that the assessment would produce identical results if the consistency of conditions is preserved, and that the test scores are free from measurement error. Tests for *validity* or *construct validity* inform the researcher whether the designed instrument provides with an adequate, meaningful and appropriate interpretation of test scores with regard to a specific setting and purpose the instrument's author has in mind. Reliability is fundamental but not necessarily sufficient step to comment on the measure's validity. (Messick, 1995; Pedhazur & Schmelkin, 1991; see also American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999, pp. 9–24).

With regard to validity, all indicators of the theoretical constructs in my thesis are according to relevant theoretical framework. *Construct validity* of these theoretical constructs is tested by the measurement models and the relationship between indicators and latent variables is captured by factor loadings. The factor loadings are high for all indicators to their latent variables, see Results section (p. 33), the chapter on "Confirmatory Factor Analysis and Reliability Estimates". *Internal validity* relates to the study in question and informs whether the cause-effect relationships are unique or whether more confounding independent variables cause the dependent variable in question. An extensive literature review (p. 7) highlighting all the relationships between the independent variable, mediators and dependent variable was provided to support the internal validity. The relationship between the independent and dependent variables is well-established empirically, and with regards to mediating relationship, the researcher claims that her sets of mediators supported theoretically and verified empirically is only one of the potential mechanisms underlying formation of SES – achievement relationship.

According to Messick (1995), *generalizability* is one of the six unified aspects within "construct validity", and it also refers to the *external validity* – the degree to which the findings would hold true for the population. Thus, before conducting analysis of educational achievement across participating countries and time, the researchers have to ensure the requirements for cross-national comparability of both manifest indices and latent variables are met (Nagengast & Marsh, 2013). The final TIMSS 2011 databases contain high-quality data, ensuring each variable and student achievement scores are internationally comparable, national adaptations are performed for all items necessary to assure *content validity* of the TIMSS instrument in each country, and sampling weights can be used for international comparisons (Neuschmidt, 2012). National centers of participating countries were responsible for checking all the science test items against their national curricula and judged them by points: 216 items equal to 233 score points (or 100% test-curriculum matching). Following this standard, Sweden identified 218 points as matching their curriculum and Ukraine – 199 (Marten et al., 2012). Thus, *content validity* matter was addressed. In this study, the *measurement invariance* (configural, metric and scalar) of four latent constructs across Sweden and Ukraine is estimated to make sure that the results can be compared between these two countries. *Criterion-related validity* tests how well one measure predicts the outcome for another measure, for example, IELTS[4] can be an effective tool in predicting further

[4] International English Language Testing System, conducted by the British Council.

students' performance in English. Criterion-related validity consists of concurrent and predictive validity, both of them are addressed within the TIMSS framework.

TIMSS 2011 survey's authors regard reliability as a crucial quality control step and use Cronbach's alpha coefficient (a) to compute test reliability at the assessment booklet level in terms of consistency within country scoring, cross-country scoring, and across assessment scoring (Foy et al., 2012). Developed by Lee Cronbach in 1951, alpha is used to reflect internal consistency of the test or a scale – to what extent the items in a test or scale measure the same underlying construct, and for high reliability its value should be close to 1: $a > 0.6$ – acceptable, $a > 0.8$ – good reliability (Cronbach, 1951; Pallant, 2011). Thus, the test reliability coefficient for the science domain in TIMSS 2011 is 0.83 in Sweden, and 0.84 in Ukraine. With regard to scales' internal consistency, the Cronbach's alpha for each scale from context questionnaires with factor loadings for each item is also computed and freely available on TIMSS and PIRLS website (Martin et al., 2012). Most of the coefficients are at an acceptable level of above 0.6 or 0.7, which will be presented for the variables included in the model in the respective chapter.

Ethical Considerations

The idea of not harming anyone on any level when collecting data, both qualitative and quantitative, is essential to the research. The secondary data provided by TIMSS 2011 for analysis does not contain any sensitive confidential information, it has ID numbers assigned to all schools and students making them completely anonymous in the published data files – pseudonymization technique. Besides, people cannot use the results to track a single sample, which is possible thanks to how the test is administered: each student only answers a section of the full test. Survey is not mandatory – students can skip answering some questions. To conclude, TIMSS data is under the protection of National Research Coordinator and permitted according to the Privacy Protection Law in Sweden.

Apart from that, it is important to remember that all empirical results have their social consequences. This aspect is especially important in the quantitative research. Since quantitative analysis is typically based on large-scale population representative data, this leads to its strength in the validity of generalization of the results. However, this may also bring some drawbacks when interpreting the empirical results. For example, certain subgroup of the population may tend to have a higher unemployment rate and lower school performance, and higher delinquency rate. Such empirical results if interpreted improperly will bring social conflicts and discrimination for that particular group. As researchers, we have to be very cautious about the possible social consequences our research results may bring when interpreted and used by others.

In the particular case of this study, significant findings about the mediation role of academic self-beliefs in shaping educational equity, which is of high priority for Sweden and many other countries, may lead to further emphasis on examining other individual student characteristics as a main player in establishing educational (in)equity. This can swap the focus from school and teacher factors to student factors, further affecting educational policies and curriculum, and, eventually, particular groups of students. However, the author cautions that the multifaceted research approach has to be applied in examining equity, and her work is only a stepping stone to investigating mediating and moderating mechanisms on other levels, i.e. teacher and school. Thus, only after studying all the underlying mechanisms behind the equity in schools, the findings shall be brought up further for consideration to education agencies, so that fairness in their decision making is preserved with regards to all players and participants of the school system.

Variables

To perform the analysis a set of student background variables, independent and mediating, was selected which was theoretically proven to potentially affect student achievement scores in science subdomains. All of the variables including science test scores are described in detail, with the list of items for two mediating factors given in Appendix 2.

The model analyzed will include exogenous or independent variables (IV) and endogenous or dependent variables (DV), as well as mediators which can serve both as IV and DV. In causal statistical modeling, exogenous variables are those which are not affected by other variables in the model of interest, however, they may have a causal effect on the endogenous variables. In turn, endogenous variables are at least partially explained by one or few variables in the hypothesized model, and have no impact on exogenous variables (Salkind, 2010). Mediating variables explain the relationship between a predictor and outcome, and are casually related to both of them, which means that they are in a casual consequence between two variables (Fritz, 2007).

Each variable within the dataset represents just a long list of values, which do not convey any useful information unless the values are summarized and the variable's distribution is examined (Muijs, 2004). There are a few common ways to describe the variable or, in other words, to perform descriptive statistics. Before doing that, a few notions will be shared with the reader. The mean is computed when measuring center of the distribution or central tendency, with the following formula: $\overline{x} = \frac{x_1 + x_2 + ... + x_n}{n}$, where the sum of observations is divided by the number of observations. Thus, mean can be also referred to as the average of a set of observations. Standard deviation (SD) measures the spread of a distribution when reporting average distance between the observations and their mean. Confidence intervals (CIs) presented in the appendices reflect within which percentage level of confidence the produced estimates are true for the population, with lower 0.5 and upper 0.5% representing 99% interval, lower 2.5 and upper 2.5% - 95% interval. For example, if the estimates lie between 95% confidence interval, it means that the researcher can be 95% sure these estimates can describe results for the target population.

Exogenous Variables

Socioeconomic Status Index

The TIMSS 2011 index for student socioeconomic status (SES) – "Home Educational Resources"

(BSBGHER, labelled as "SES" in the model for the ease of reader's understanding[5]) – was computed based on student responses to three multiple-choice questions: parents' highest education level (BSDGEDUP), number of home study supports (BSDG05S) and number of books at home (BSBG04) (Martin et al., 2012). The three indicators to SES with their mean and standard deviation are represented in the *table 3* below, while all the items included within each indicator can be found in the *Appendix 1*.

Table 3. Descriptive Statistics of Home Educational Resources scale.

		Sweden				**Ukraine**			
Indicator	Variable Label	Min	Mean	Max	SD	Min	Mean	Max	SD
SES	Index of home educational resources (continuous variable)	4.95	11.31	14.02	1.61	4.95	10.65	14.02	1.47

Gender

Gender is a dichotomous variable which was dummy coded for the model proposed in this study, with values "0" for girls and "1" for boys in both the datasets for Sweden and Ukraine. It is used in the model as a control variable in order to control for the effects gender might have on students' confidence or value on science and in turn, on their science achievement. In Sweden, 48% of girls and 52% of boys participated in the test, while in Ukraine the percentage of boys- and girls-participants was equal, see *Table 2.*

Ethnicity

The second factor placed in the model as a control variable is ETHNIC, which contains information on the student's ethnic background. As in our model we are focused solely on the students' non-cognitive mediators of SES impact on science performance, irrespective of gender, ethnicity or possible external influences, it is important to exclude such potential variables which might affect this mediating relationship. Specifically, it is more important to account for ethnicity factor in Sweden, as it has a heterogeneous population, consequently making ethnic background a more significant factor in the interplay among SES - science self-concepts - science achievement variables.

ETHNIC is a dichotomous dummy-coded variable with two choices: "Yes" with value "1", and "No" with value "0" in Sweden and Ukraine data. These were answers to the question on whether students were born in the country where they are taking TIMSS test. Thus, those with "0" had immigrant background and those with "1" were nationals of the country, although the possibility of having immigrant parents was not covered by this question. In Ukraine, the percentage of born in the country who participated in the test was 97%, in Sweden, it totaled to 91%.

Parental Support

The third control variable which represents possible external influences mentioned above is PSupport – parental support. It is a latent construct which consists of four items related to the frequency of parental

[5] As the researcher mentioned in the Definitions of Key Terms section, when interpreting results it should be kept in mind that Home Educational Resources index is only part of socioeconomic status, and other measures of it are available, uni- and multi-dimensional.

involvement in their child's school performance matters, and assessed by students on the 4-category Likert scale: 1 - "Every day or almost every day", 2 - "Once or twice a week", 3 - "Once or twice a month", 4 - "Never or almost never" (*table 4*).

Following standard deviation values, we can establish that there is more variability in the level of parental support in Swedish families compared to Ukraine. Besides, the mean shows us that parents in Sweden supervise their children's school progress less often, mostly once or twice a week (mean value around "2"), while in Ukraine the frequency of supervision varies between at least few times a week and every day. In both cases, parental support still is a significant factor, thus it is important to control for its effects, if reliable results are to be concluded regarding the mediating effects of students' individual motivational constructs.

Table 4. Descriptive statistics of Parental Support manifest variables.

		Sweden		**Ukraine**	
Indicator	Variable Label	Mean	SD	Mean	SD
PSuppor1	How often do your parents ask what you learned in school?	2.03	0.91	1.46	0.77
PSuppor2	How often do you talk about schoolwork with your parents at home?	1.87	0.86	1.57	0.79
PSuppor3	How often do your parents make sure that you set aside time for your homework?	2.06	1.02	1.25	0.61
PSuppor4	How often do your parents check if you do your homework?	1.95	1.02	1.93	1.05

Endogenous Variables

Students Confident in Science

This is a complex scale (*see Appendix 2*) which consists of nine statements/items, answered by eighth graders separately for each science subdomain in Sweden and Ukraine on a four-point Likert scale: "agree a lot", "agree a little", "disagree a little", "disagree a lot". The questions were of the type: "I learn things quickly in biology/chemistry/physics/earth science", "Biology/chemistry/physics/earth science is harder for me than any other subject/ than for my classmates" etc. The author has referred to the scale as complex because it is multifaceted and encompasses a variety of factors that build students' confidence in their abilities in science subdomains – self-concept, peer- and subject-comparison, teacher encouragement and anxiety (Bandura, 1997). This is a mediating variable.

Four continuous manifest indices in total were used for the Students Confident in Science latent construct (SCICONF), which descriptive statistics and Cronbach's alpha reliability coefficient are presented in the *table 5* below. The indicator names were changed for the ease of the analysis and reader's understanding. Cronbach's alpha is above 0.8 for all the continuous indices which is very good and means that internal consistency of the scales is high. The Swedish students' level of confidence is somewhat higher than of Ukrainian students, with less standard deviation from the mean (central tendency of replies on scale items), although in both cases the mean corresponds to "Somewhat Confident" level as per TIMSS 2011 Context Questionnaire Scale Details (Martin et al., 2012).

Table 5. Descriptive Statistics and Reliability coefficient for Students Confident in Science Subdomains.

		Sweden			Ukraine		
Indicator	Variable Label	Cronbach's *a*	Mean	SD	Cronbach's *a*	Mean	SD
ConfBIO	Students Confident in Biology	0.88	9.76	1.73	0.87	10.09	1.97
ConfCHE	Students Confident in Chemistry	0.87	10.19	1.67	0.91	9.71	2.17
ConfPHY	Students Confident in Physics	0.86	10.14	1.67	0.88	9.88	1.96
ConfEAR	Students Confident in Earth Science	0.87	10.15	1.82	0.87	9.73	1.9

Students Value Science

Second mediator, utility or extrinsic value of science, is measured in TIMSS 2011 by "Students Value Science" six-item scale representing students' perceived importance of broad science for their future studies or career prospects, or for their life in general, assessed on a same four-point Likert scale (*please refer to Appendix 2*). As both countries have science taught as separate four subjects, the latent construct SCIVALUE was made the same way as the construct for students' confidence in science. It employed four manifest continuous indices reflecting students' value of biology, chemistry, physics and earth science.

Cronbach's alpha reliability coefficient is equally high for each continuous variable. Students' value of science subdomains varies between 9.4 and 9.81 in Sweden, and 9.86 – 9.98 in Ukraine, which is slightly less than confidence but lies within "Somewhat Value" category. Standard deviation from the mean is also smaller which means that eighth graders' replies are less varied and are closer to the central tendency (*Table 6*).

Table 6. Descriptive Statistics and Reliability coefficient for Students' Value of Science Subdomains.

		Sweden			Ukraine		
Indicator	Variable Label	Cronbach's *a*	Mean	SD	Cronbach's *a*	Mean	SD
ValueBIO	Students Value Biology	0.87	9.4	1.6	0.86	9.98	1.87
ValueCHE	Students Value Chemistry	0.89	9.5	1.56	0.88	9.9	1.89
ValuePHY	Students Value Physics	0.89	9.41	1.62	0.87	9.91	1.81
ValueEAR	Students Value Earth Science	0.86	9.81	1.6	0.85	9.86	1.77

Students Like Learning Science

Interest or intrinsic value of science is represented by "Students Like Learning Science" five-item index reflecting the extent to which students like physics/chemistry/biology/earth science and enjoy learning science, assessed on a four-point Likert scale. The latent construct SCILIKE comprises of four manifest continuous indices reflecting students liking to learn biology, chemistry, physics and earth science. As seen from the *table 7*, eighth graders in Ukraine showed more interest in science than in Sweden with one-point difference. This is a third mediator the researcher intends to use in her model.

Table 7. Descriptive Statistics and Reliability coefficient for Students' Like Learning Science Subdomains.

		Sweden			**Ukraine**		
Indicator	Variable Label	Cronbach's *a*	Mean	SD	Cronbach's *a*	Mean	SD
LikeBIO	Students Like Biology	0.89	9.19	1.79	0.86	10.74	1.85
LikeCHE	Students Like Chemistry	0.89	9.55	1.8	0.89	10.34	2.09
LikePHY	Students Like Physics	0.89	9.32	1.75	0.88	10.52	1.97
LikeEAR	Students Like Earth Science	0.88	9.48	1.79	0.87	10.34	1.91

Multicollinearity

Before including all the mediating latent measures in the analysis, it is important to test that they and their indicators are not highly correlated between each other, because if they do, the results of certain statistical tests within SEM method may be biased (Khine, 2013; Kline 2011). When any pair of variables has correlation higher than r = .85, one of the two variables should be excluded from further analysis.

Three mediators represented as latent constructs: SCICONF, SCIVALUE and SCILIKE were checked for the level of correlation in Mplus. The correlation between SCICONF and SCILIKE was very high (r = 0.87) which basically means that these two latent constructs measure almost the same thing. SCICONF with SCIVALUE: r = 0.49; SCIVALUE with SCILIKE: r = 0.70. Thus, after the test the decision was taken to exclude the SCILIKE construct which represented a potential threat for the good model-fit due to its high correlation coefficient.

Missing Data

The examining of observed variables (predictor, control variables and mediators) have been performed, leaving us to one more step – summing up the number of missing values on variables or missing responses. All missing values are given the value of -99 in SPSS and are accounted while doing modelling in Mplus software which excludes cases with missing on all variables and gives total number of such cases. Nevertheless, it is useful to know the percentage of missing information in order to do appropriate conclusions.

In Sweden, the sum of missing values for ConfBIO-ConfEAR and ValueBIO-ValueEAR varies between 12 and 25%. However, Mplus uses inbuilt multiple imputation with EM algorithm to handle the missing data and maximum likelihood estimation with robust standard errors (MLR) can achieve unbiased parameter estimates. Confirmatory factor approach (CFA) will be applied to them in order to check for factor loadings and standard errors. Few more techniques like Imputation and Bootstrapping will help to overcome bias when running the structural model, which will be discussed later (Field, 2017). The amount of missing data in other variables in the current analysis is typically around 2% in both Sweden and Ukraine.

In Ukraine, there is very small percentage of missing responses on all the variables, for more details please refer to the *table 8*.

Table 8. Number and missing values for all the manifest or observed variables.

	Sweden		Ukraine	
Indicator	Number	Missing %	Number	Missing %
SES	5459	2	3366	0.4
Gender	5502	1.3	3375	0.1
ETHNIC	5488	1.5	3366	0.4
PSuppor1	5452	2.2	3341	1.1
PSuppor2	5447	2.3	3338	1.2
PSuppor3	5442	2.4	3343	1
PSuppor4	5451	2.2	3339	1.2
ConfBIO	4893	12.2	3349	0.9
ConfCHE	4567	18.1	3345	1
ConfPHY	4565	18.1	3333	1.3
ConfEAR	4195	24.7	3353	0.7
ValueBIO	4901	12.1	3344	1
ValueCHE	4546	18.4	3341	1.1
ValuePHY	4548	18.4	3328	1.5
ValueEAR	4178	25	3348	0.9
Total	5573		3378	

Student Achievement in Science

As we know, TIMSS and other large-scale assessments' design implies that estimation of students' proficiency in a particular domain is done with only a small subset of responded test items from each student, based on which inferences are made about the population achievement where sampling error applies. Because of a big percentage of missing information, such way of computing test results will have some level of uncertainty – a measurement error which is added up to a sampling error (Wu, 2005; von Davier et al., 2009). In order to express this measurement uncertainty, five multiple scores or imputations are generated for each student to improve the accuracy of estimate and error (Laukaityte & Wiberg, 2017). These imputations are called plausible values (PV) and represent student's ability estimate, with the degree of variation between five scores expressing the level of error. We cannot regard PV as testing scores of an individual student, because they are randomly drawn from the estimated distribution of student ability parameters which is called as posterior distribution (Wu, 2005). This means that although two students might give same responses on test items, the random drawing of plausible values from the distribution of their abilities can result in different achievement scores for both of them. Eventually, ILSAs objective is to estimate not individual characteristics but to describe population performance and that is what PVs are good for.

In Sweden and Ukraine, five plausible values are calculated for each subject within the domain of science (biology, chemistry, physics and earth science), and also for the science domain in general.

Descriptive statistics of the first plausible value for the science domain and all subdomains are presented in *table 9* as an example. We should not consider science scores from 1st PV as final, because TIMSS presents the average of five plausible values to report *International Results on Science* (Martin et al., 2012). Nevertheless, variability between plausible values is small and it is sufficient to analyze one PV in order to have a clear picture of results' distribution. As we can see, Sweden and Ukraine have almost same distribution of science achievement scores': the difference between the scores of students at the 90th and 10th percentiles in Sweden – 207, and in Ukraine – 208, and is bigger for the subdomains. The average of science 1st PV is 508 for both countries.

Table 9. Science Achievement in Sweden and Ukraine, TIMSS 2011 (1st Plausible Value)

SWEDEN							
	Mean	Min	Max	Standard Deviation	10th Percentile	90th Percentile	90th – 10th
Science (general)	507.91	204.07	761.09	81.30	400.50	607.33	206.83
Biology	509.63	190.89	792.10	84.87	398.25	613.04	214.79
Chemistry	501.40	177.61	794.81	85.66	387.81	606.83	219.02
Physics	494.75	122.69	830.04	90.34	376.03	604.51	228.48
Earth Science	516.51	170.38	872.32	91.93	393.01	630.96	237.95
UKRAINE							
	Mean	Min	Max	Standard Deviation	10th Percentile	90th Percentile	90th – 10th
Science (general)	508.22	226.98	764.11	81.18	401.19	609.53	208.34
Biology	499.90	223.35	767.49	79.52	396.39	599.57	203.18
Chemistry	518.90	109.32	784.92	91.66	394.95	634.01	239.06
Physics	510.90	144.49	775.37	91.28	389.27	624.74	235.47
Earth Science	501.17	131.62	835.44	96.62	375.46	620.60	245.14

With the purpose to maximize the reliability of results and minimize the measurement error we will run our analysis five times using each time a different set of PVs from four science subdomains which will generate the average of five parameter estimates. Such procedure is called a multiple imputation model type and was suggested by Rutkowski, Gonzales, Joncas, and von Davier (2010). It is implemented in Mplus software (Muthén & Muthén, 1998–2012) and our structural model estimates will be generated using TYPE=IMPUTATION command for five datasets. Means for biology (PVBIO), chemistry (PVCHE), physics (PVPHY) and earth science (PVEAR) are presented in *table 10*, and confidence interval (CI) distribution can be referred to in the *Appendix 3*. When checking CI table please keep in mind that all values were divided by 100 as it simplifies Mplus estimation procedure.

Table 10. Means for four science subdomains based on the average of five plausible values for each.

Country	PVBIO	PVCHE	PVPHY	PVEAR
Sweden	512.50	502.30	498.00	519.70
Ukraine	492.40	511.50	502.70	494.70

For the reader to be able to infer from these score values, it is important to mention that according to TIMSS 2011 International Results report, four international benchmarks are set to identify what students know and can do based on their science scores (Martin et al., 2012). Students with score points of 400-475 had low proficiency level, 475-550 – intermediate proficiency, 550-625 – high proficiency. In 2011, Swedish and Ukrainian students performed slightly above the 500 scale centerpoint with their average of 509 and 501 accordingly. We can see, however some variability in their abilities in different science subdomains.

Structural Equation Modeling as Analytical Method

To test my theoretical model with TIMSS data, Structural Equation Modeling (SEM) is used in this study. SEM modeling technique was developed in the 1970s and owned more and more popularity in the social and behavioral sciences. Path analysis is the most commonly used modeling method where direct and indirect effects (i.e., mediation effects) can be examined, and the theoretical mechanism among different constructs in focus can be tested (Khine, 2013).

SEM is a multivariate statistical analysis technique which takes on a confirmatory (hypothesis-testing) approach in examining the relationships between multiple observed and unobserved variables while providing explicit estimates of error variance parameters, as well as direct and indirect effects of variables being studied. Other multivariate techniques cannot be employed for assessing or correcting for measurement error, and no easy alternative methods are available for examining *indirect effects*. Besides, in SEM both independent and dependent variables can be defined as latent constructs. These features and the fact that SEM has been widely and effectively used in studying cause-effect mechanisms in large-scale assessment surveys, have resulted in the researcher's choice of this method (Khine, 2013; Byrne, 2012).

To distinguish observed (manifest) variables and unobserved (latent) variables, an example of Parental Support construct is given (*Figure 3*). Manifest variables are psuppor1-psuppor4 graphically presented in squares (or rectangles), each corresponding to a response related to how often parents get involved in their children's school work. These four manifest variables are indicators to the underlying latent variable PSupport presented in a circle (or, sometimes, ellipse), and the choice of indicators is theoretically driven.

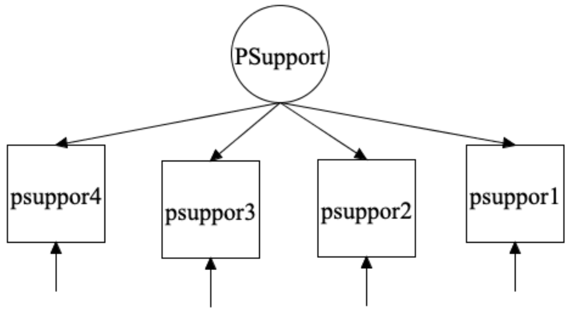


Figure 3. A measurement model of parental support latent variable.

Structural equation modeling comprises of measurement and structural models, which sometimes represent a two-step approach to SEM emphasizing on estimating whether the measurement model holds, examined by model fit indices (see Model evaluation below), before proceeding with testing structural relationships (Jöreskog, & Sörbom, 2003). *Figure 3* represents a *measurement or CFA model*, which assesses the fit of the observed variables to the latent variable, estimating factor loadings of each indicator on the latent variable, as well as computing measurement errors for observed variables and residuals for the underlying factor, displayed below (Khine, 2013; Byrne, 2012). Both measurement errors and residuals are presented as residuals in Mplus, however we will refer to them as error and residual in the Results chapter.

 - measurement error – to what extent is manifest variable adequate in measuring the latent variable

 - residual error – the extent to which the predicted value of endogenous latent variable is in error

The *structural model* or the second step of SEM defines the nature and magnitude of relationships between the latent variables (factors) and estimates the regression coefficients of factors on manifest variables (Khine, 2013). In this thesis a *recursive structural model* which specifies a one-way causal relationship is used (Byrne, 2012). There are two types of effects in SEM: direct and indirect. In the case of direct effects, single-direction arrows are used depicting the relations among latent constructs, however, directionality does not imply causality. Indirect effects, which are of main interest to this study, reflect how the relationship between exogenous (independent) and endogenous (dependent) variables is influenced or mediated by one or more latent variables. Thereby, our structural model will be a *Multiple Mediator Model*, where mediating variables can be defined as:

> "…intermediate in the causal sequence relating an independent variable to a dependent variable. That is, the independent variable causes the mediating variable which then causes the dependent variable" (MacKinnan, 2008).

Most commonly, SEM models are tested during five steps: model specification, identification, estimation, evaluation, and modification (Khine, 2013), and the importance of following them applies to this study too.

Model specification

Figure 4 represents the first model specification step, where relations among variables are hypothesized based on the proposed conceptual framework. This is a structural model with multiple mediators where measurement models for latent constructs are incorporated, including both the direct effects between the exogenous manifest variable *SES, gender*, *ethnic* and *PSupport* –and endogenous latent variable *TotScie* (outcome), as well as the indirect effects from *SES, gender*, *ethnic* and *PSupport* via the latent mediators *SciConf* and *SciValue* on the outcome variable *TotScie*. To be more precise for model specification in Mplus, *TotScie* is regressed on *SES* – direct effect, TotScie regressed on *SES* via *SciConf* – specific indirect effect, TotScie regressed on *SES* via *SciValue* – specific indirect effect, TotScie ON SES via both SciConf and SciValue – total indirect effect (presented as "TotScie IND SES" in Mplus); same applies to the regression on control variables. Double-headed arrows depict covariances or hypothesized correlations between the pairs of independent variables or, in our case, between two mediators as well. Such graphical representation of a model reflecting hypothesized relationships among all the variables is called *path diagram*.

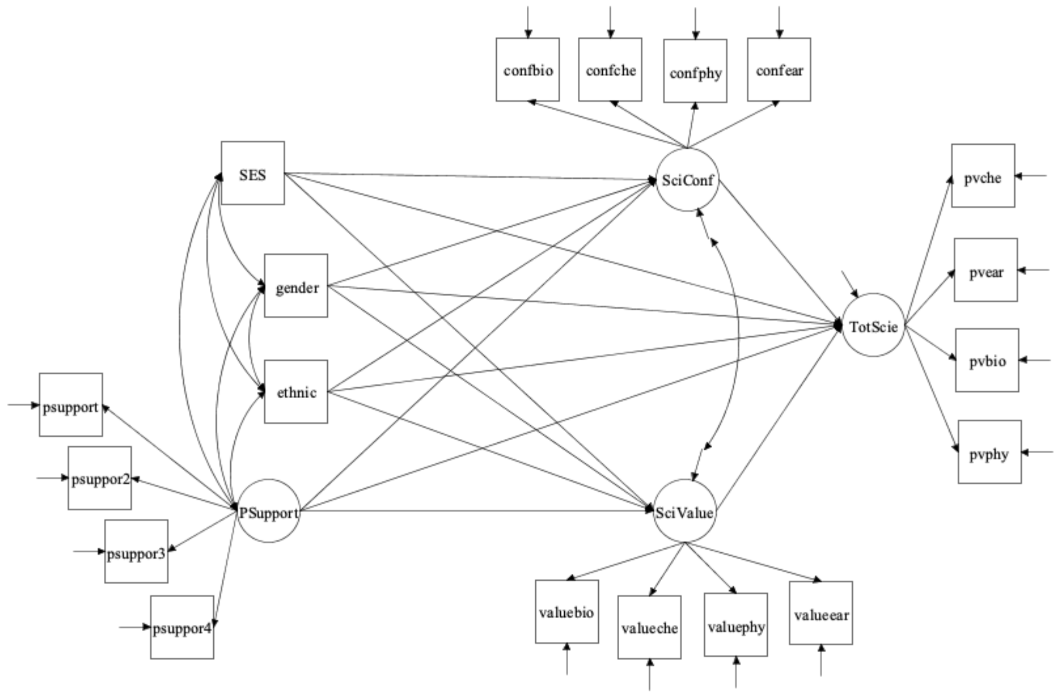


Figure 4. A hypothesized structural model with mediating relationships.

There are three types of paths to be specified within the diagram: directional effects – between observed variables (called factor loadings) and latent constructs or among different latent variables (called path coefficients); variances and covariances. The relations among the variables called as paths or parameters can be set to *fixed*, *free* or *constrained parameters* (Khine, 2013). However, Mplus software is somewhat easy to use for novices in statistical analysis as it specifies values for the parameters by default (Byrne, 2012): (1) the first factor loading in each set of indicator variables is set to 1.0 in the measurement portion of our SEM model, (2) residual variances of and covariances among independent latent variables including mediators (SciConf, SciValue and PSupport) are freely estimated, and (c) error variances associated with each dependent manifest variable are freely estimated, reflected by one-headed arrow pointing at squares.

Our model is formulated for both Sweden and Ukraine. Because it comprises of four CFA models and structural model, latent variables have to be well-measured by their indicators. Thus, factor loadings will be estimated for the four latent constructs and their measurement invariance will be tested across Sweden and Ukraine to enable the comparison between the results in two countries. Besides, a two-level analysis has been applied first in order to check for the intraclass correlation coefficient (ICC), to be presented. Just to remind that the author's focus is on student-level individual characteristics mediating family background (socioeconomic status) where only one-level model with complex option is sufficient to control for standard error estimates, and initial application of two-level model is important to support that decision.

Model Identification

In the second step of model identification we proceed with analyzing whether the model is testable. This is done by transposing the "known" information – variance/covariance matrix (or correlation matrix and standard deviations) of the measured variables (the data), – into the "unknown" – the structural parameters of the model under study (e.g., factor loadings, factor correlations, measurement errors) (Khine, 2013; Byrne 2012). There are three criteria or types of the model identification. It is not possible to identify a model if the number of parameters to be estimated exceeds the "known" information, and that is called as "under-identified" type of model. Thus, for an acceptable "just-identification" type, the number of data points (variances and covariances of the observed variables) or "knowns" must be exactly the same as the number of parameters to be estimated – "unknowns". We call the model as "overidentified" when "known" information exceeds the "unknown" resulting in generating positive *degrees of freedom* (*df*). This allows for the model rejection and proves its scientific use. Therefore, in identifying the model researchers always look for the "over-identification" criterion. In the current analysis, all measurement models are over-identified, and so is the final path model presented above in Figure 2 (i.e., with 209 known parameters and 75 unknown parameters). For more elaborate discussions on the topic please refer to Bollen and Davis (2009), where they give a detailed mathematical explanation to some of the rules of identification.

Model Estimation

Once the model is identified, it is necessary to ensure that the population parameters are estimated with the minimum difference between "the observed (sample) variance/covariance matrix and the model-implied (model-predicted) variance/covariance matrix" (Khine, 2013). Maximum likelihood (ML) estimator has been a default algorithm for use with missing data patterns in Mplus and many other SEM softwares for quite a while now. Just like multiple imputation, it is a robust missing data handling procedure and enables the model to be converged at a higher speed (Enders, 2010). As per Khine's summarized description, ML "is an iterative technique, which means that an initially posited value is subsequently updated through calculation. The iteration continues until the best values are attained. When this occurs, the model is said to have converged" (2013). The main assumption for and ML and

MLR (robust maximum likelihood) is that the scale of observed variables is continuous, and not categorical, while the sample size should be at least 100-150 participants (Byrne, 2013). The bigger is the sample size, the higher is sensitivity analysis of ML method to the differences among the data. MLR estimator also enhance robustness against non-normal distribution of the data.

Intraclass Correlation Coefficient

As described before, TIMSS has a hierarchically structured data with students clustered on the class level and nested within school clusters. As a rule, at least a two-level analysis is applied to such data, but, with the focus on individual characteristics of students, the researcher initially planned for a single-level model. Nevertheless, estimating intra-class correlation coefficients (ICC ρ) is an important step before proceeding with model evaluation. By computing ICC, we can find out to what extent the variations in student's performance can be attributed to his or her belonging to different groups, such as schools or classrooms, and make a conclusion about the equity (Yang Hansen, 2014). If mean differences in the level of performance are large between the different groups, the ICC becomes large, which can cause type I error – the rejection of the null-hypothesis – if no appropriate analysis technique is employed.

Therefore, when ICC is computed, a decision can be taken on the modelling technique, and appropriate cluster-size for parameter estimates' scaling. School-size cluster is usually selected for the analysis as it gives more variability among students' achievement within school samples than within classrooms. In order to prove it and provide the rationale for selecting school- and not class-size cluster, as well as to give reasoning for selecting a single-level or multilevel technique, the intra-class correlations (ICC) measuring between-school and between-class differences for student achievement scores in science subdomains, their socioeconomic status (SES) and responses to two scales of our interest in Sweden and Ukraine were estimated (see *Input 1 in Appendix 6*).

The ICC coefficients for school cluster-size scaling varied between 0.16 and 0.19 for the science subdomains in Sweden, which means that 16 to 19% of variance in student science performance is explained by the effects of schools. For Ukraine, the between-school variance coefficients were larger, varying between 0.18 and 0.25. Thus, 18 to 25% of variance in science test scores in TIMSS 2011 is explained by the school effects. With regard to SES, 14% of students within a school sample belong to the same socioeconomic background in Sweden, amounting to 24% – in Ukraine. This tells us that in Ukraine the distribution of students between schools depends on their socioeconomic status more than in Sweden. The ICC coefficients for between-classroom variance were even larger, confirming the appropriateness of selecting school cluster for the purpose of scaling (for the rest of estimates please refer to *Table 11*).

When estimating ICC, it is imperative that coefficients should be below 0.1 in order to avoid biased standard error estimates and type I error. In case of ICC coefficients larger than 0.1, a multilevel analysis is usually required (Hox, Maas, & Brinkhuis, 2010). However, after applying a two-level model to both the Sweden and Ukraine data, the intraclass correlation values for the manifest variables of two our latent constructs of interest – Student Confident in Science and Students Value Science – are much lower than 0.1 in Sweden and lower or around 0.1 in Ukraine. Moreover, when comparing the parameter estimates achieved from the two-level analysis to those from single-level model with complex option to adjusting standard error estimates due to the hierarchical data structure, no different conclusions can be reached. This finding and the fact that the author of this study is primarily focused at the individual-level non-cognitive characteristics prove that it is sufficient at the current stage of research to use the command COMPLEX in Mplus software to control for cluster effects and run a single-level model.

Table 11. Estimated School- and Class-level ICCs of Science Achievement, Students Confident in Science and Value Science scales, and SES for Grade 8 in Sweden and Ukraine.

	Country	Sweden		Ukraine	
		Class	**School**	**Class**	**School**
	Clusters / Average cluster size	266 / 36	153 / 21	162 / 21	148 / 23
Science subdomains	Biology	0.22	0.19	0.25	0.24
	Chemistry	0.21	0.17	0.27	0.25
	Physics	0.21	0.18	0.26	0.25
	Earth Science	0.18	0.16	0.19	0.18
Students Confident in Science	Biology	0.06	0.04	0.11	0.11
	Chemistry	0.06	0.05	0.1	0.09
	Physics	0.05	0.03	0.1	0.1
	Earth Science	0.05	0.04	0.09	0.09
Students Value Science	Biology	0.04	0.03	0.12	0.12
	Chemistry	0.05	0.03	0.12	0.11
	Physics	0.03	0.02	0.12	0.11
	Earth Science	0.03	0.02	0.11	0.11
SES		0.15	0.14	0.24	0.24

Model Evaluation and Model Fit

The fourth step of SEM is testing for how well the model with its variance/covariance matrix fits the variance/covariance matrix of the data used for the study (Byrne, 2012; Khine, 2013). As mentioned before, the goal is to minimize the difference between these two matrices. Commonly, a statistically nonsignificant chi-square ($\chi 2$) value is used to indicate a good fit – that the difference between variances and covariances of the observed data and theoretical model is nonsignificant. Chi-square falls into the absolute-fit indices category which measures how well the specified model reproduces the data. However, chi-square test is not useful with large samples produced by such large-scale assessment surveys like TIMSS and others (Khine, 2013). As a result, many other model fit indices have been developed and tested by the researchers in order to decide on accepting or rejecting the proposed model, which are well-summarized e.g. by Kline (2011).

There are three categories of fit indices in total: absolute fit, comparative fit and parsimonious fit. The most used and referred to fit indices are divided between these three categories, they are also the ones presented by default in Mplus (Khine, 2013; In'nami, & Koizumi, 2011; Bandalos & Finney, 2010):

CFI (comparative fit index), TLI (Tucker-Lewis index) – comparative fit indices; RMSEA (root mean square error of approximation), and SRMR (standardized root mean square residual) – absolute model fit indices. Chi-square mean ($\chi 2$) and degrees of freedom (*df*) are presented in model fit table by default, but reporting *p*-value which will always be significant because of a big sample size is irrelevant. In turn, RMSEA corrects chi-square's tendency to reject the models with large sample size or number of variables as it has high sensitivity to the number of estimated parameters (Khine, 2013; Hooper, Coughlan, & Mullen, 2008). RMSEA can be also used as an indicator of model parsimony, which is assessed as the ratio of *df* used by the specified model to the total *df* available.

CFI and TLI comparative fit indices assess whether the specified model is better than a competing baseline model, one that assumes that all observed variables are uncorrelated, with values ranging between 0 and 1. In a good model fit CFI and TLI ≥.95, values for acceptable model fit should be ≥.90. SRMR and RMSEA indicate the extent or error generating from the hypothesized model's estimation. Lower RMSEA and SRMR (≤.06) values indicate good model fit, values ≤.08 – acceptable fit, values ≥.10 – unacceptable fit. In some cases, researchers report RMSEA ≤.08 as a good model fit.

Measurement Invariance which examines e.g. whether the factor loadings are the same across groups (in our case, across Sweden and Ukraine) is also tested using model fit indices, which will be discussed in the Results chapter.

Model Modification

After checking for ICC among main factors of interest and across class- and school-level clusters, it is decided to run a single-level model with complex option for the evaluation in Mplus. TYPE=COMPLEX and ESTIMATOR=MLR functions in ANALYSIS command in combination with CLUSTER=IDSCHOOL and WEIGHT=HOUWGT (house weight) in VARIABLE command will adjust for cluster effects, non-independence of observations and unequal selection probability, and account for multiple missing data patterns, while house weight is used for the purpose of computing estimations relevant for the individual student level. Bias-corrected bootstrap method (BOOTSTRAP in Mplus) with confidence intervals is also applied to the structural model instead of MLR estimator to see if more reliable mediated effects can be produced (MacKinnon, 2008). There are certain limitations to one-level complex model where only individual level variables can be studied. In case of incorporating more variables at a teacher- or school-level which will be proposed in *Conclusions*, a multilevel modeling technique is required for providing reliable estimates.

Results

In this chapter we will start with confirmatory factor analysis (CFA) for four latent indices used in our specified structural model, which corresponds to a two-step approach to SEM referred to in our introductory chapter to this method. Then, measurement invariance of latent variables is estimated which is a common procedure in cross-national comparisons of the models with latent variables (Desa, 2014;

Wendt, Kasper, & Trendtel, 2017). As a final step, the evaluation of the specified model will be presented which will answer the research questions, providing rationale for rejecting or accepting hypotheses.

Confirmatory Factor Analysis and Reliability Estimates

According to Jöreskog and Sörbom (2003), a two-step approach to SEM involving estimating the measurement model before proceeding with testing structural relationships is essential. Only having confirmed that the measurement model holds (expressed in model-fit indices), the parameter estimates from the structural model can be considered as reliable. Thereby, the model fit information is represented below for all four latent variables (*table 12*), as well as a separate table with factor loadings, standard error (S. E.) estimates and residuals to reflect their content validity and reliability for the purpose of evaluation within the specified structural model (*table 13*).

Table 12. Model fit indices for four measurement models representing four factors.

Latent Indices	Sweden				Ukraine			
	$\chi2$ (df)	CFI	RMSEA	SRMR	$\chi2$ (df)	CFI	RMSEA	SRMR
Students Confident in Science (SciConf)	30.5 (2)	0.98	0.05	0.02	27.3 (2)	0.98	0.06	0.02
Students Value Science (SciValue)	11.2 (2)	1	0.03	0.01	74.7 (2)	0.96	0.1	0.03
Parental Support (PSupport)*	358.2 (2)	0.92	0.18	0.06	56.7 (2)	0.9	0.09	0.03
Science Achievement (TotScie) # of computations: 5	111 (2)	0.99	0.09	0.005	65.8 (2)	1	0.09	0.01
Good fit Acceptable Threshold		≥.95 ≥.90	≤ 0.08 ≤ 0.10	≤ 0.06 ≤ 0.10		≥.95 ≥.90	≤ 0.08 ≤ 0.10	≤ 0.06 ≤ 0.10

SciConf and SciValue latent variables are made of continuous indices representing students' confidence and value in four science subdomains, and these are reliable estimates generated by TIMSS 2011 team. TotScie (total science achievement) average estimates are computed with IMPUTATION technique using different sets of plausible values which enables to considerably reduce measurement errors. PSupport (parental support) comprises of four categorical variables, thus WLSMV estimator (mean- and variance-adjusted weighted least squares) was used in its CFA model to compare of estimates produced with MLR, which is used primarily for continuous variables, will be same. The WLSMV is a robust estimator for modelling categorical data as it does not assume normally distributed variables (Brown, 2014). After comparing the CFA model with both WLSMV and MLR, the estimates were exactly the same, therefore, MLR estimator was chosen to represent results for all the measurement models.

Table 13. Factor loadings, standard errors and residuals of four latent variables.

Latent Indices		Sweden			Ukraine		
		Standardized results					
		Estimates	Standard Error* (S.E.)	Residual Variances	Estimates	Standard Error* (S.E.)	Residual Variances
Students Confident in Science (SciConf)	CONFBIO	0.71	0.02	0.49	0.64	0.02	0.59
	CONFCHE	0.88	0.01	0.22	0.65	0.03	0.58
	CONFPHY	0.80	0.01	0.36	0.66	0.02	0.56
	CONFEAR	0.49	0.02	0.76	0.71	0.02	0.49
SCICONF with SCIVALUE: r = 0.48 (correlation)					SCICONF with SCIVALUE: r = 0.46 (correlation)		
Students Value Science (SciValue)	VALUEBIO	0.73	0.01	0.46	0.74	0.02	0.45
	VALUECHE	0.90	0.01	0.20	0.77	0.02	0.42
	VALUEPHY	0.86	0.01	0.26	0.67	0.02	0.55
	VALUEEAR	0.58	0.02	0.67	0.74	0.02	0.46
Parental Support (PSupport)	PSuppor1	0.53	0.02	0.72	0.58	0.03	0.67
	PSuppor2	0.52	0.02	0.73	0.46	0.04	0.79
	PSuppor3	0.83	0.01	0.30	0.48	0.04	0.77
	PSuppor4	0.80	0.01	0.36	0.50	0.04	0.74
Science Achievement (TotScie)	PVCHE	0.93	0.01	0.14	0.93	0.01	0.14
	PVEAR	0.93	0.01	0.14	0.92	0.01	0.15
	PVBIO	0.95	0.01	0.09	0.94	0.01	0.12
	PVPHY	0.93	0.01	0.13	0.94	0.01	0.12

- *According to Muthén and Muthén (2002), standard error bias should be < 0.10 (10%) for any parameter in the model, while S.E. bias for the parameter for which power is of particular interest should be < 0.05 (5%). Factor loadings ≥ 0.4 are acceptable, ≥ 0.6 – good.

In general, model fit indices in Table 13 showed that the measurement models fit the data well. Chi-square statistics is large, that is due to the fact that chi-square is sensitive to large sample size and tends to be significant (see Table 3). It is therefore important to examine other model fit indices when sample size is large, as the case here.CFI, RMSEA and SRMR are all under the cutoff value indicating good model fit. After analyzing all four latent variables, the conclusion can be made that all of CFA models fit well to the data, and factor loadings are high for all of them with very small standard errors. RMSEA is 0.18 for PSupport variable in Sweden, however, it is acceptable based on other model fit indices. The measurement invariance test for four factors will be applied now to check for their comparability.

Measurement Invariance

Before proceeding with comparing results in Sweden and Ukraine, we have to make sure that the latent constructs (factors) created for both countries are invariant. Factors comprise of many scales/ manifest indicators built on questionnaire items, and this statistical step is required to see if same underlying construct is being measured across groups or time. Basically, the test for measurement invariance reflects whether our measures mean the same thing for participants belonging to different groups.

Thus, CFA models of four factors – SciConf, SciValue, PSupport and TotScie (based on 1st PVs) – are fit into Sweden's and Ukraine's sample in order to see if scales of latent variables are invariant on configural, metric and scalar levels. Categorical CFA modelling requires additional parameters to increase its robustness which is beyond the scope of this thesis, thereby model-fit indices are projected to reflect lower level of acceptability for the categorical latent construct PSupport (Desa, 2014). In Mplus, MODEL= Configural Metric Scalar; function is added to ANALYSIS command to run the test and model-fit indices (CFI, RMSEA and SRMR) are used to describe the invariance (see *Table 14*). Metric invariance is the minimum requirement for the parameter estimates to be compared across two countries. There are three most commonly used levels of invariance from the basic and less restricted to the most restricted (Rutkowski & Rutkowski, 2013): a test for *configural invariance* estimates whether the same number of indicators is loaded per latent variable across groups; *metric invariance* tests whether the factor loadings are same across groups; *scalar invariance* reflects whether scale's item averages are the same across groups.

Table 14. Measurement invariance of Students Confident in Science continuous latent construct.

Invariance Level		CFI	RMSEA	SRMR	Δ CFI*	Δ RMSEA
	Good fit	≥.95	≤ 0.08	≤ 0.06		
	Acceptable threshold	≥.90	≤ 0.10	≤ 0.10		
Configural		0.98	0.06	0.02	-	-
Metric		0.96	0.06	0.09	0.022	**0.008**
Scalar		0.91	0.08	0.12	0.049	*0.015*

CFI – Comparative Fit Index
RMSEA – Root Mean Square Error Of Approximation
SRMR – Standardized Root Mean Square Residual

* According to Chen (2007), model is invariant across two groups if absolute change in CFI (Δ CFI) ≤ .01, and absolute change in RMSEA (Δ RMSEA) ≤ .01. Values in bold – absolute invariance, italicized values – acceptable change.

As we can see from the table, Students Confident in Science scale is invariant on all parameters at a configural and metric levels. At a scalar level, it is also at the acceptable invariance on two parameters (CFI and RMSEA). The change in RMSEA between metric and scalar levels is acceptable, however, the change in CFI is not. Measurement invariance tables for Students Value Science, Parental Support and Science achievement latent index based on 1st plausible values from 4 subdomains can be referred to in *Appendix 4*. The minimum requirement is achieved across all the scales allowing for cross-country comparison.

Model Fit of the Final Structural Model for Sweden and Ukraine

Just like model fit was performed with measurement models, it is applied now for the final structural model fit into Sweden's and Ukraine's sample to confirm its workability and reliability of the parameter estimates to be analyzed. The model was run in Mplus twice for both countries, first time – using MLR as estimator, and second – using BOOTSTRAP function in ANALYSIS command. BOOTSTRAP is a resampling method that enables to replicate estimations as many times as specified by the researcher (e.g. 5-10000 times), each time with a slightly different sets of samples randomly picked by the software. It is not that essential with big samples but useful in case of focusing on indirect effects, in particular. To explain the method better let us refer to MacKinnan: "Assume that you have a sample of size N. The bootstrap method consists of randomly sampling with replacement from the original N observations so that a new sample of N observations is obtained, which is the first resample (or the first bootstrap sample)" (2008). Only RMSEA and SRMR indices are estimated in Mplus with this command (*table 15*).

Table 15. Model fit of the structural model with two mediators.

Model-fit Indices			Sweden		Ukraine	
	Good fit	Acceptable threshold	MLR	BOOTSTRAP	MLR	BOOTSTRAP
χ2 (df)			3993 (134)		2871 (134)	
CFI	≥.95	≥.90	0.91		0.86	
RMSEA	≤ 0.08	≤ 0.10	0.07	0.04	0.08	0.04
SRMR	≤ 0.06	≤ 0.10	0.05	0.05	0.05	0.05

df – degrees of freedom
χ2 – Chi-square mean
CFI – Comparative Fit Index
RMSEA – Root Mean Square Error Of Approximation
SRMR – Standardized Root Mean Square Residual

*After rounding up the values, TLI for Ukraine's model is somewhat lower 0.90 threshold, however, RMSEA and SRMR meet the good fit standard, and CFI – acceptable threshold standard, which is enough for the model to be considered reliable.

Although RMSEA value was lower with BOOTSTRAP function, after comparing confidence intervals (CI) of outputs produced in Mplus with MLR and BOOTSTRAP estimators, there was a non-significant difference of .001 – .002 in the lower and upper 0.5% of CI, averages remained the same. Therefore, the diagrams and estimates will be presented with MLR estimator only, confidence intervals tables for total, total indirect, specific indirect, and direct effects are given in *Appendix 5* for Sweden and Ukraine, which will be discussed further.

Reporting Results for Sweden and Ukraine

Now that all the required steps before evaluating the structural multiple mediator model have been completed, it is time to present and discuss the results in order to answer our research questions and test hypotheses. We are going to look at both the models run with Sweden's and Ukraine's dataset, because significant relationships are same for the constructs of interest with some differences for control variables (*figure 5, figure 6*). Some of the most powerful effects will be discussed in the next chapter comparing Sweden and Ukraine with all the estimates presented in a *table 16.*

Only significant paths have been left in the path diagram to give a clearer picture to the reader, and we can see that all the paths of our interest are significant with a p-value < 0.005. However, for the purpose of reporting more meaningful results, I will only report effects as significant if $p < 0.001$. There is also such notion as effect size (z-effect) which is used in examining the power of the association between variables and is derived by the formula z = estimate/S.E. (parameter estimate divided by standard error).

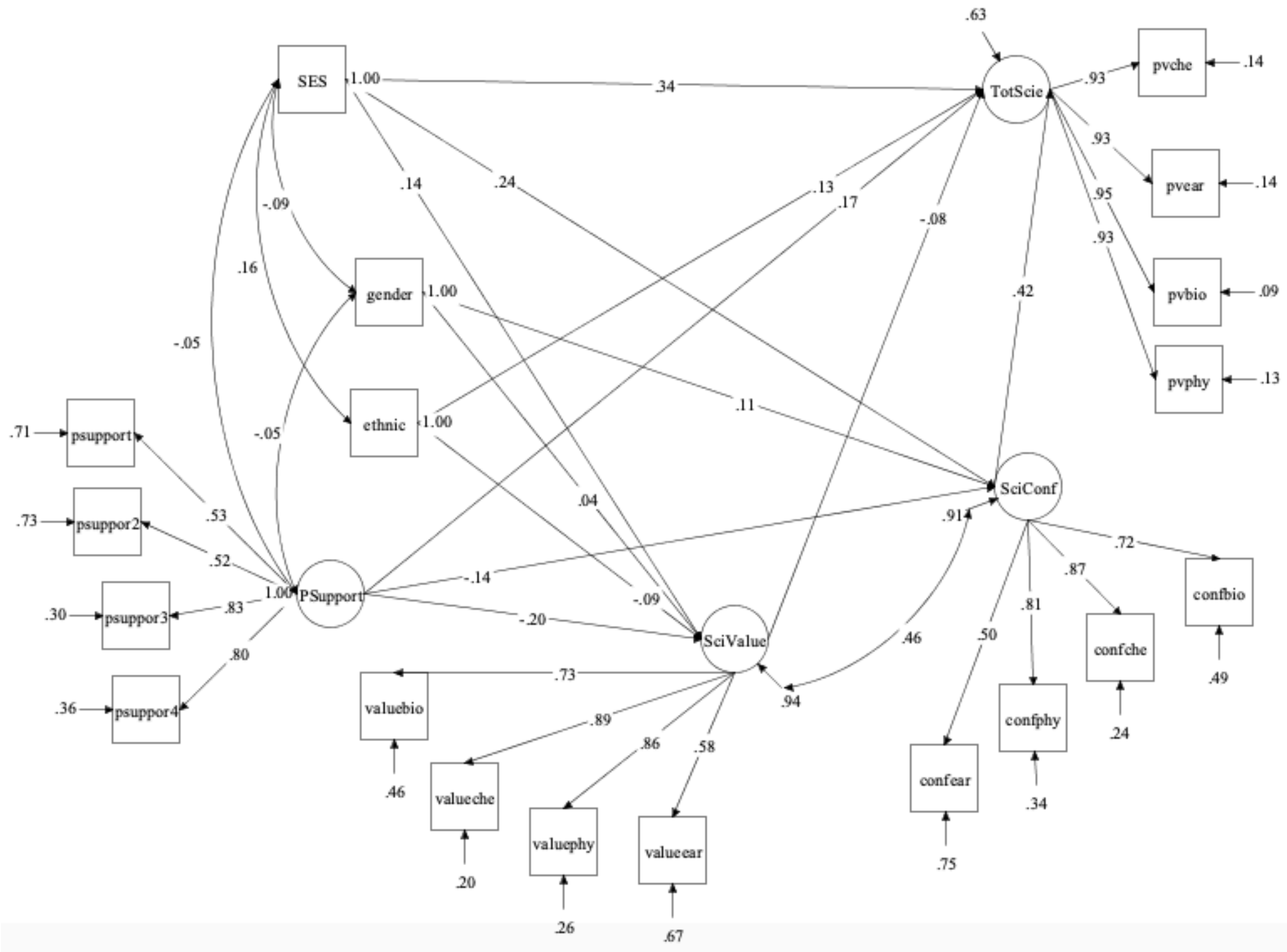


Figure 5. Structural model for Sweden with significant paths.

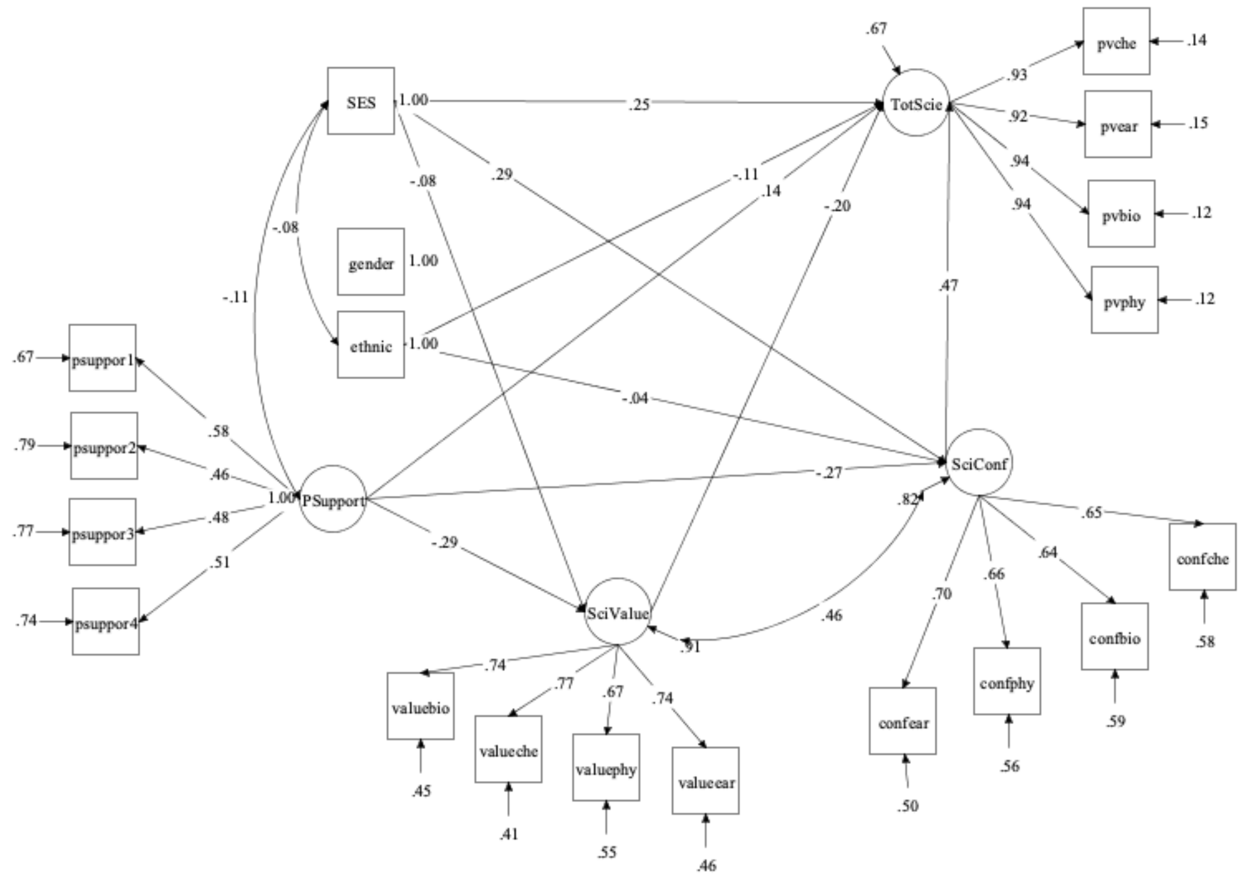


Figure 6. Structural model for Ukraine with significant paths.

Hypothesis 1. Judging from the model, we conclude that students' socioeconomic status (SES) is related to science self-confidence (SciConf) and utility (SciValue), in turn, science-self-confidence and utility factors are casually related to science achievement (TotScie). Rejecting this null-hypothesis is essential to prove the validity of the proposed theoretical model (for coefficients refer to *table 16*).

Hypothesis 2. Students' science self-confidence indirect effect from SES to TotScie in Sweden is 0.101, which is calculated by multiplying path coefficients of SES to SciConf and SciConf to TotScie: 0.24 x 0.42 = 0.101. In Ukraine, the indirect effect is higher with value of 0.138. Both mediating effects are significant, and we can accept the second alternative hypothesis H2 that science self-confidence partially compensates for inequity in science achievement resulting from students' background level in Sweden and Ukraine.

Hypothesis 3. SES has a significant effect on students' value of science (SciValue), however, SciValue has a small negative effect on TotScie which means that students valuing science surprisingly perform slightly lower. The indirect effect of SciValue is again calculated by multiplying two path coefficients: SES to SciValue and SciValue to TotScie. It is very small and non-significant in both countries with our $p < 0.001$, however, it reduces the indirect effect of SciConf in Sweden and adds up to SciConf indirect effect in Ukraine. SciConf still holds significant in both cases. To conclude, we accept our $H3_0$ that the indirect effect of science utility is not significant to partially compensate for inequitable educational outcomes as a consequence of students' socioeconomic status. Same applies to both countries.

Hypothesis 4. Following the discussion above we reject the fourth null-hypothesis that the mediating effect of science self-confidence will NOT hold when simultaneously considering the indirect effect of utility, because SES – SciConf – TotScie relationship holds significant even after accounting for SciValue in both Ukraine and Sweden. However, it is non-significant for SES – SciValue – TotScie.

Hypothesis 5: We partly reject the fifth null-hypothesis, as only the mediating effects of self-confidence (SciConf) on SES – science achievement relationship has remained significant after controlling for students' gender, ethnicity and parental support variables.

All hypotheses have been tested, and a more elaborative explanation and comparison of findings between Ukraine and Sweden will now be provided.

Comparing Significant Findings in Sweden and Ukraine

Starting with educational equity and deducting from our models, in Sweden 34% of variance in science achievement can be attributed to students' socioeconomic status with two science self-beliefs mediating ¼ of SES – science achievement relationship (total indirect effect – 0.091). Meanwhile in Ukraine more equitable educational outcomes are observed with only 25% of variance resulting from SES, with more than half of the association between socioeconomic status and science performance mediated by science self-confidence and value of science (0.154). Although in both cases science value contributes only a small effect, non-significant at $p < 0.001$. All the estimates are presented in *table 16*.

Self-confidence or self-concept in science has a significant positive indirect effect for both countries, which alone mediates one third of SES effect on TotScie in Sweden and half of the SES effect in Ukraine. This is an important finding to prove that the contribution of individual factors should not be overlooked when assessing educational (in)equity. It can be further suggested that the interplay of a whole range of factors on individual, class and school-level, both mediating and moderating have to be studied in order to conclude on the question of equity in any country.

In contrast to science self-confidence, student's value in science is negatively associated with science achievement, which basically means the more value students assign to science the worse they perform. This is a bit contradictive but interesting observation as it is shown in both countries with higher negative impact accounting to 20% of variance in Ukraine, and only 8% in Sweden. Socioeconomic status also affects utility negatively in Ukraine (8%) meaning that the value of science reduces with higher SES, which is opposite to Sweden where higher socioeconomic status increases value of science for the future with 14% of variance in SciValue contributed by SES.

There is an interesting relationship among Parental Support (PSupport) – science self-beliefs (SciConf and SciValue) – TotScie constructs, which is observed for both countries. In Ukraine, as discussed before, parents get involved in their children's school work twice more often than in Sweden, and there is twice higher negative indirect effect in Ukraine among PSupport – SciConf – TotScie relationship. All in all, parental support has a significant negative impact on both mediators and positive – on science achievement, which is a reversed causality example. It means that children with low science performance and low confidence need more parental support and involvement, and this requirement is higher for students in Ukraine, because low parental involvement causes higher level of under-confidence in science among Ukrainian students. This dependency on parental support is a cultural phenomenon, but it also be a signal that teachers do not get enough time and resources to cater for everyone's individual needs, with as many as 35-40 students given to one teacher to handle in big towns and cities. Thus, parental involvement is important.

With regard to remaining two control variables, gender effects are non-significant in Ukraine. In Sweden, although gender does not have a significant direct effect on science achievement, its effect is mediated by self-confidence which is of small significance and this mediated relationship contributes to 5% of variance in TotScie for boys. The explanation to why gender plays a non-significant role in science achievement in Ukraine can lie in different factors. The equality in treatment irrespective of gender is of course guaranteed by the Ukrainian Education Law, but there are no specific measures introduced to propagate the gender equality apart from the fact that teachers attempt to maintain equal treatment in the classroom. This can be especially emphasized as 93% of teachers who participated in the TIMSS 2011 in Ukraine were women, and it could be suggested that girls get more privileged treatment. It rather can be considered as another historically or culturally-adopted phenomenon, where women always did all kinds and types of work of any level of difficulty and complexity, whether physical or intellectual, at par with men. This includes science, medical, machinery and other fields. Thus, the strive for success is not determined by gender, but simply by the personal abilities. In Sweden, there is a very strong message about the gender equality and no direct effect of gender on science achievement was found. Nevertheless, there is still a small but significant indirect effect via science self-confidence, which means that boys can translate their self-confidence into a better performance.

Ethnicity variable (the fact whether a student was born in the country where assessment was taken or not; Yes = 1, No = 0) has only significant direct effect on science performance in two countries, albeit, non-significant when mediated by individual self-beliefs. It should be noted that we cannot check for the immigrant background with this variable, as a person can be born in the country and still be of a different ethnicity, it is rather a fact confirming that a student feels belonging in the country and speaks its national language. When checking for effect size, ethnicity in Sweden influences science outcome twice as much as in Ukraine. In Sweden, the SES is positively correlated with ethnicity, which means that those born in the country have higher SES and higher achievement in science. In Ukraine, it seems that the ones who were born outside the country have higher SES and achievement which is not as surprising as it may seem. As a rule, people move to Ukraine only from the post-Soviet countries, they already speak Russian and children easily learn Ukrainian at schools, therefore, the language is not a hurdle. The common reason for moving to Ukraine from nearby countries is usually for those who are planning on setting a business and already have a good financial stability. Also, it could be military families who are stationed in different countries and have good income. Thus, children who were born outside the country will have higher SES and somewhat higher achievement.

Table 16. Total, total indirect, specific indirect, and direct effects estimated in the structural model.

	SWEDEN				UKRAINE			
% of factor variance explained by model	9% SCICONF explained 7% SCIVALUE explained 37% TOTSCIE explained				18% SCICONF explained 9% SCIVALUE explained 33% TOTSCIE explained			
	Effects				**Effects**			
	Total (z-effect)*	Direct (z-effect)	Indirect (z-effect)	p-value[1]	Total (z-effect)	Direct (z-effect)	Indirect (z-effect)	p-value
Effects from SES to TOTSCIE								
TOTSCIE to SES	0.428 (24.826)	0.337 (19.164)	0.091 (9.904)	0.000	0.402 (14.808)	0.249 (8.958)	0.154 (9.872)	0.000
TOTSCIE SCICONF SES			0.101 (9.877)	0.000			0.138 (9.222)	0.000
TOTSCIE SCIVALUE SES			-0.011 (-3.018)	0.003			0.016 (2.933)	0.003
Effects from PSUPPORT to TOTSCIE								
TOTSCIE to PSUPPORT	0.130 (7.096)	0.173 (9.731)	-0.043 (-5.002)	0.000	0.067 (2.010)	0.136 (3.940)	-0.069 (-4.109)	Direct and indirect – 0.000 Total – 0.044
TOTSCIE SCICONF PSUPPORT			-0.058 (-6.610)	0.000			-0.126 (-7.341)	0.000
TOTSCIE SCIVALUE PSUPPORT			0.015 (3.197)	0.001			0.058 (5.770)	0.000
Effects from GENDER to TOTSCIE								
TOTSCIE to GENDER	0.028 (1.963)	-0.016 (-1.171)	0.044 (6.777)	Total – 0.050 Direct – 0.242 Indirect – 0.000	0.055 (2.383)	0.036 (1.554)	0.018 (1.720)	Total – 0.017 Direct – 0.120 Indirect – 0.085
TOTSCIE SCICONF GENDER			0.047 (6.804)	0.000			0.013 (1.067)	0.286
TOTSCIE SCIVALUE GENDER			-0.003 (-1.815)	0.070			0.006 (1.182)	0.237
Effects from ETHNIC to TOTSCIE								
TOTSCIE to ETHNIC	0.141 (8.589)	0.126 (7.699)	0.015 (2.368)	Total and direct – 0.000 Indirect – 0.018	-0.132 (-5.369)	-0.108 (-4.174)	-0.024 (-2.759)	Total – 0.000 Direct – 0.000 Indirect – 0.006

TOTSCIE SCICONF ETHNIC		0.009 (1.338)	0.181		-0.019 (-2.404)	0.016
TOTSCIE SCIVALUE ETHNIC		0.007 (2.659)	0.008		-0.005 (-1.066)	0.286

- *z-effect is derived via Estimate/S.E., and it reflects the effect size or the power of the relationship between variables.
- [1]p-value is significant at < 0.001

Conclusions and Discussions

The aim of this study was to examine how socioeconomic status and science achievement relationship is mediated by the expectancy-value model proposed by Eccles & Wigfield (1983) based on which science-related self-concept, intrinsic and extrinsic values were matched to the measures present in TIMSS 2011 for assessing eighth graders' attitudes towards science subdomains. The rationale for studying mediators of SES influences was to emphasize once again that there is much more to educational (in)equity than just positive association between SES and student performance. And this association starts with individual prerequisites of students which underlying mechanisms in shaping differences in educational outcomes are to be accounted for at all levels of study: individual, class, school, society. It can be attempted to reduce the gaps in student achievement results or maintain them at the same level, however, once genetic predisposition has been considered (for details see Johnson, McGue, & Iacono, 2007; Spinath et al., 2008), it is in fact undesirable for educational system to equalize student outcomes. Only the system where a student can realize his or her full potential can be called equitable which consequently will lead to the seemingly unequal distribution of achievement scores.

Our findings contribute to the existing prior work in the area of studying mediating mechanisms behind socioeconomic status – academic achievement association, and partly confirm those of Kriegbaum, & Spinath (2016), and Steinmayr, Dinger, & Spinath (2012) that academic self-concept significantly mediates academic achievement. Interest measure had to be dropped out because it was empirically indistinguishable from science self-concept construct, albeit its high correlation with science self-concept (r = 0.87) signals that conclusion about partial mediation can be extended to interest in science. On the other hand, domain-related utility value did not prove to be a significant mediator of the relationship between home educational resources index and science achievement. Instead, science value had a small negative indirect effect in Sweden and two-sided negative effects in Ukraine, with only 7-9% of science value explained by the hypothesized model.

One of the possible explanations for this can be Eccles' EVT model itself (1983; Wigfield & Eccles, 2000) where students value of a particular domain is steered by their self-concept of being competent in that domain, thereby science self-concept predicts utility value and it is not a two-headed relationship as suggested by the model. Academic self-concept in turn at many times depends on the individual's interpretation of contextual factors (frames-of-reference) which leads us to one of the paradoxical effect of Big-Fish-Little-Pond (BFLPE theory) (Marsch, Parker & Pekrun, 2018), and Information Distortion Model (IDM) suggested by Parker et al. (2018) who referred to BFLP effects and rational action theory (Goldthorpe, 2007) in developing their model of how social classes affect students' values and self-beliefs via ability stratification which takes place at the moment of school choice. This is especially applicable to Sweden with its highly decentralized school system, and schools representing a market where students exercise their free choice as consumers.

The fundament for BFLPE theory is that students compare their educational outcomes with those of their classmates and use this peer-comparison as the basis for their academic self-concept (Marsh, 1987; Marsh et al., 2015; Parker et al., 2014). Therefore, taking as an example two students with same abilities but from different social background, the one with lower socioeconomic status will go to school with mixed- or low-achievement, and the one with high SES will attend a school where average achievement is high. Consequently, academic self-concept will exacerbate of the one with disadvantaged background and drop of the one attending a school with high achievement. At this point Parker's et al. IDM model comes into picture (2018) which empirically proves that such tendency may only seem as equalizing force where higher academic self-concept of student with low SES will enable him to achieve same level of outcome as his high-SES-peer. Nevertheless, in reality the more stratified education system is the more restrictions a student from disadvantaged background will have to translate his or her high academic self-concept into high educational outcomes, one of such restrictions being lack of resources and the very educational structure itself. According to Parker et al. (2016), this inability of transforming higher academic self-beliefs into corresponding level of achievement results in the weaker relationship between academic self-concept, task values and performance, as well as negative association between SES and students' aspirations.

This theory can be an explanation to what we see in our model results. Only 9% of science self-concept and 7% of utility value are explained by the model in Sweden which has an ability stratified educational system, whereas Ukraine's model has explained 18% of academic self-concept and 9% of value. The education system in Ukraine is mostly centralized with a very small share of private or specialized schools. The mediating effect of self-concept in Sweden has less power than that in Ukraine, and the total mediating effect is reduced by a small negative influence of science value in Sweden. For explaining a small negative effect of value – achievement relationship, contextual factors have to be examined, some of which were discussed above. Directional relationships between SES and value, and value – science achievement are negative in Ukraine, however, the cumulative mediating effect transforms this negative association into positive.

Parental involvement control variable is another factor of interest, which mediating role in SES – achievement association was recently studied by Guo et al. (2018). In our case, however, it is represented as a variable mediated by students' individual non-cognitive factors which effects on science achievement are controlled for in the study. A reversed causality effect was found which means that low performers and low-confidence students need more parental support and involvement. It contradicts the findings represented in Guo et al.'s work (2018) which stresses out that home monitoring played a negative role among girls after controlling for SES and parental expectation. In Ukraine, this requirement for parental support is twice as in Sweden.

Gender or ethnicity differences were not focus of our research. It can be mentioned that gender did not play any role in the Ukraine case. In the case of Sweden, gender had no direct effect on science achievement, but gender influence was mediated through confidence in science which implied that boys have higher self-confidence in science and in turn out-perform girls in science achievement, totaling to 5% of confidence in science translating for boys into better performance. Ethnicity had no significant indirect effect but a small power direct effect on science outcome in both countries.

After analyzing the results, it can be claimed that the presented research furthers the dialogue about necessary examination of student-level factors mediating the formation of (in)equity in education systems. It also supports the claim that academic self-concept, referred to sometimes in recent research as self-confidence (Sheldrake, 2016), is one of the best non-cognitive predictors of educational outcomes and a significant mediator of socioeconomic impact on achievement even without considering interest and utility value. In order to discuss policy implications of these findings though, contextual factors moderating the mediating relationship have to be added on multiple levels.

The considerations for further research are discussed below.

Limitations and Further Research

First limitation to name is that this thesis is a cross-sectional and not longitudinal study of students' non-cognitive mediating factors, and their effects indeed are better to be studied over a period of time, in our case, it is suggested that few TIMSS surveys should be analyzed. All three scales of students' academic self-concept or self-confidence, interest and value of science were improved in TIMSS 2015 including more items assessing corresponding construct. Therefore, a same snapshot of mediating effects of psychological factors can be repeated for TIMSS 2015 and including more countries for the comparison.

A single-level model represented our focus on students' individual factors and their home background. However, finding significant mediation of SES – achievement relationship attributed to some of the student non-cognitive factors like academic self-concept is not enough to conclude on the level of educational equity. Multilevel structural equation modeling can be applied to include contextual factors on teacher- and school-level that will moderate student's self-concept, interest and value, and might help to explain the rest of 63-67% of science achievement left unexplained by our model. Nevertheless, our research just set a cornerstone for reference, especially related to Sweden, when conducting further multilevel examination of mechanisms shaping educational (in)equity, setting a reminder that student psychology factors in evaluating equity are just as important as macrolevel factors being studied. And if extended multilevel research in Sweden will prove suggested here influences of ability stratification schooling system which weakens the relationship between student non-cognitive factors and achievement (Parker et al., 2018), observed in our model in contrast with Ukraine, it will be of significant finding for policy discussions.

Some minor limitations are that in order to generalize findings about non-cognitive characteristics mediating educational equity, same study has to be repeated for both mathematics and reading domains assessed in TIMSS in Sweden and Ukraine. Kriegbaum, & Spinath (2016) did examine longitudinal mediating effects of cognitive and non-cognitive factors related to mathematics domain but based on PISA conducted in Germany, while the sample size in Steinmayr, Dinger, & Spinath (2012) work was relatively small.

Last but not the least factor that the researcher considers as the limitation is the way the index of socioeconomic status presented in TIMSS 2011 was made. SES index is constructed on the basis of three indicators: number of books at home, parents' highest level of education and home study supports. There is no doubt that a fundamental work is undertaken to assess the response items of TIMSS background questionnaire with getting feedback from National Research Coordinators (NRCs) by the IEA, selecting particular items for a field test, and revising field-test instruments for the final data collection (Rutkowski & Rutkowski, 2010). However, as Rutkowski & Rutkowski pointed out (2010), there might be issues with missing data treatment (in Sweden, 44% of students did not reply about their parents' level of education), poor scale reliability for particular countries (internal scale consistency was only $a = 0.38$ in Sweden, and 0.45 in Ukraine which is below usually accepted reliability coefficient of 0.5-0.6), and occasional inaccuracy of test participants in answering the questions. Besides, it will be useful to separate socioeconomic status into indicators and attempt to find differences in mediating effects related to, for example, father's or mother's highest education level or family cultural resources, like number of books at home, only.

References

1. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, DC: AERA.
2. Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 93-114). New York: Routledge.
3. Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
4. Bandura, A. (2012). On the Functional Properties of Perceived Self-Efficacy Revisited. *Journal of Management*, *38* (1), pp. 9-44.
5. Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, *72*, 187–206.
6. Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
7. Beach, D., Dovemark, M., Schwartz, A., & Öhrn, E. (2013). Complexities and Contradictions of Educational Inclusion: A Meta-Ethnographic Analysis. *Nordic Studies in Education*, *33* (4), p. 254-268.
8. Bollen, K. A., & Davis, W. R. (2009). Two rules of identification for structural equation models. *Structural Equation Modeling*, *16*, 523–536.
9. Brown, T. (2006). *Confirmatory factor analysis for applied research*. London: The Guilford Press.
10. Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research*, Second Edition. New York: Guilford Publications.
11. Byrne, B. M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. New York: Routledge.
12. Chao, R. K., & Willms, J. D. (2002). The effects of parenting practices on children's outcomes. In J. D. Willms (Ed.), *Vulnerable children: Findings from Canada's National Longitudinal Survey of Children and Youth,* pp. 149–166. Edmonton, AB: University of Alberta Press.
13. Chen, F. F. (2007), Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 3, pp. 464-504.
14. Chohan, B. I., & Khan, R. M. (2010). Impact of parental support on the academic performance and self-concept of the student. *Journal of Research and Reflections in Education*, *4* (1), 14–26.
15. Coleman, J. (1968). *The concept of equality of educational opportunity*. Harvard Educational Review, *38* (1), 7–22.

16. Craven, R. G., & Marsh, H. W. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology*, *25* (2), 104-118.
17. Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16:* 297-334.
18. Desa, D. (2014). *Evaluating Measurement Invariance of TALIS 2013 Complex Scales: Comparison between Continuous and Categorical Multiple-Group Confirmatory Factor Analyses.* Paris: OECD Publishing.
19. Eccles (Parsons), J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation,* pp. 75–146. San Francisco, CA: Freeman.
20. Eccles, J. (2009) Who Am I and What Am I Going to Do with My Life? Personal and Collective Identities as Motivators of Action, *Educational Psychologist*, *44* (2), 78-89.
21. Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53* (1), 109–132.
22. Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.
23. Espinoza, O. (2007). Solving the equity-equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research*, *49*, 343–363.
24. Field, A. (2017). *Discovering statistics using IBM SPSS statistics.* SAGE Publications.
25. Filiatreau, S. (2011). Ukraine's Participation in the Bologna Process: Has it Resulted in More Transparency in Ukrainian Higher Education Institutions? *International Research and Review: Journal of Phi Beta Delta Honor Society of International Scholars, 1* (1).
26. Foy, P. (2012). Estimating Standard Errors for the TIMSS and PIRLS 2011 Achievement Scales. In M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
27. Foy, P., Arora, A., & Stanco, G.M. (Eds.). (2013). *TIMSS 2011 international database and user guide.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
28. Foy, P., Arora, A., & Stanco, G.M. (Eds.). (2013). TIMSS 2011 Sampling Stratification Information, Supplement 4. In *TIMSS 2011 international database and user guide.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
29. Foy, P., Martin, M. O., Mullis, I. V. S., & Stanco, G. (2012). Reviewing the TIMSS and PIRLS 2011 Achievement Item Statistics. In M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
30. Fritz, D. S. (2007). Mediation Analysis. *Annual review of psychology, 58* (1): 593-614.
31. Gasa, V., Pitsoane, E., Molepo, L., & Lethole, P. (2018): The effect of families' socioeconomic status on the self-concept development of learners, *Early Child Development and Care*.

32. Goldthorpe, J. H. (2007). *On sociology* (2nd ed.). Stanford, CA: Stanford University Press.
33. Gu, L. & Kristoffersson, M. (2015). Swedish Lower Secondary School Teachers' Perceptions and Experiences Regarding Homework. *Universal Journal of Educational Research, 3* (4), pp. 296-305.
34. Guo, G., & Harris, K. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, *37* (4), 431–447.
35. Guo, X., Lv, B., Zhou, H., Liu, C., Liu, J., Jiang, K. and Luo, L. (2018). Gender Differences in How Family Income and Parental Education Relate to Reading Achievement in China: The Mediating Role of Parental Expectation and Parental Involvement. *Frontiers in Psychology*, *9*: 783.
36. Gustafsson, J.-E., & Yang Hansen, K. (2018). Changes in the Impact of Family Education on Student Educational Achievement in Sweden 1988–2014, *Scandinavian Journal of Educational Research*, *62* (5), 719-736.
37. Gustafsson, J.-E., Nilsen, T., & Yang Hansen, K. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*, *57*, 16–30.
38. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Press.
39. Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, *52*, 123–142.
40. Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
41. Helgøy, I., Homme, A. (2006). Policy Tools and Institutional Change: Comparing education policies in Norway, Sweden and England. *Journal of Public Policy*, *26* (2), pp.141-165.
42. Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles*, 2.
43. Hox, J. J. & Bechger, T. M. (1998). An Introduction to Structural Equation Modeling. *Family Science Review*, *11*, 354-373.
44. Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications.* New York: Routledge.
45. Hox, J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, *64*, 157–170.
46. Hsin, A., Xie, Y. (2017) Life-course changes in the mediation of cognitive and non-cognitive skills for parental effects on children's academic achievement. *Social Science Research*, *63*, 150-165.
47. In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly*, *8*, 250-276.

48. International Association for the Evaluation of Educational Achievement (IEA). (2012). *International database analyzer (Version 3.0)*. Hamburg, Germany: IEA Data Processing and Research Center.
49. Johnson, W., McGue, M., & Iacono, W. G. (2007). How parents influence school grades: Hints from a sample of adoptive and biological families. *Learning and Individual Differences*, *17*, 201–219.
50. Joncas, M., & Foy, P. (2012). Sample design in TIMSS and PIRLS. In M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
51. Jöreskog, K. G., & Sörbom, D. (2003). LISREL (Version 8.54) [Computer software]. Chicago: Scientific Software.
52. Khine, M. S. (2013). *Application of Structural Equation Modeling in Educational Research and Practice*. Rotterdam. The Netherlands: Sense.
53. Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.
54. Krapohl, E., & Plomin, R. (2015). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, 1–7.
55. Kriegbaum, K., & Spinath, B. (2016) Explaining Social Disparities in Mathematical Achievement: The Role of Motivation. *European Journal of Personality*, *30,* 45–63. Pub. in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/per.2042.
56. Laukaityte, I. & Wiberg, M. (2017) Using plausible values in secondary analysis in large-scale assessments, *Communications in Statistics - Theory and Methods*, *46* (22), 11341-11357.
57. Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal*, *41* (3), 423–447.
58. MacKinnon D. P. (2008). *Introduction to statistical mediation analysis.* New York, NY: Taylor & Francis.
59. Marks, G. N. (2013). *Education, social background and cognitive ability: The decline of the social*. Routledge.
60. Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors? *Educational Research and Evaluation*, *12* (2), 105–128.
61. Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*, 280–295.
62. Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. A. (2015). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology*, *107*, 258–271.

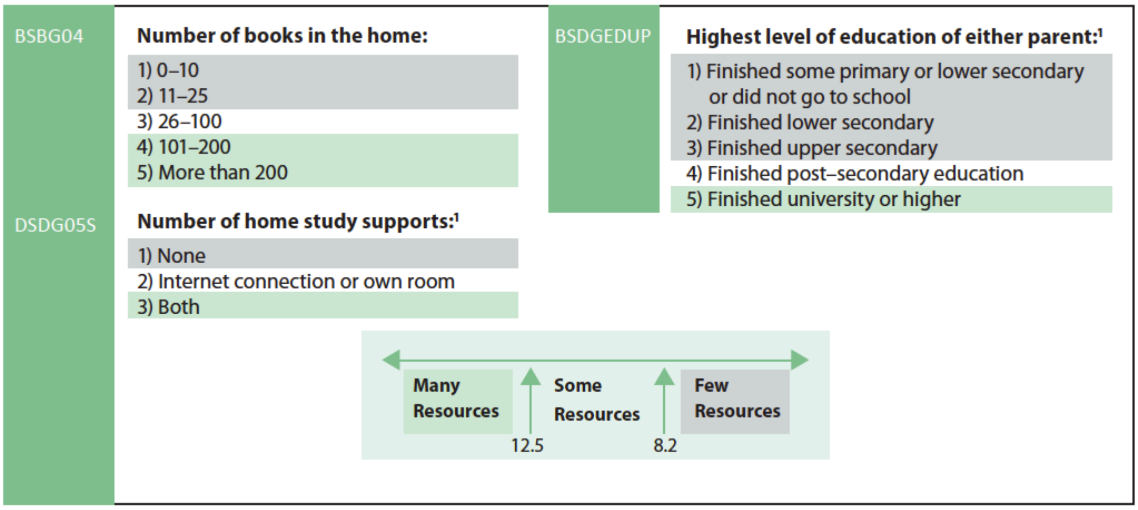
63. Marsh, H. W., Parker, P. D., & Pekrun, R. (2018). Three Paradoxical Effects on Academic Self-Concept Across Countries, Schools, and Students: Frame-of-Reference as a Unifying Theoretical Explanation. *European Psychologist*. Advance online publication.
64. Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*, 397–416.
65. Marsh, H., Craven, R., & Debus, R. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: a developmental perspective. *Merrill-Palmer Quarterly*, *45* (4), 567–601.
66. Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco G. M. (2012). *TIMSS 2011 International Results in Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
67. Martin, M. O., Mullis, I. V., Foy, P., & Arora, A. (2012). Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales. In M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
68. Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396.
69. Messick, S. (1995). Validity of psychological assessment: Validity of inferences from a person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741-749.
70. Mood, C., Jonsson, J. O., Bihagen, E. (2012). Socioeconomic persistence across generations: the role of cognitive and non-cognitive processes. In: Ermisch, John, J€antti, Markus, Smeeding, Timothy (Eds.), *From Parents to Children. The Intergenerational Transmission of Advantage*. Russell Sage Foundation, New York.
71. Muijs, D. (2004). *Doing Quantitative Research in Education with SPSS*. SAGE Publications Ltd.
72. Muijs, D. (2012). Methodological change in educational effectiveness research. In C. P. Chapman, P. Armstrong, A. Harris, D. R. Muijs, D. Reynolds, & P. Sammons (Eds.), School effectiveness and improvement research, policy and practice: Challenging the orthodoxy, pp. 58–66. Abingdon, UK: Routledge.
73. Mullis, I. V. S. (2002). *Background Questions in TIMSS and PIRLS: An Overview*.
74. Mullis, I.V.S., & Martin, M. O., Editors (2009). *TIMSS 2011 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.
75. Mullis, I. V. S., Martin, M. O., Minnich, C. A., Stanco, G. M., Arora, A., Centurino, V. A. S., & Castle, C. E. (2012). *TIMSS 2011 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*, *2*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

76. Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus users guide* (7th ed.). Los Angeles, CA: Authors.
77. Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*, 599-620.
78. Nagengast, B., & Marsh, H. W. (2013). Motivation and engagement in science around the globe: testing measurement invariance with multigroup structural equation models across 57 countries using PISA 2006. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis*, Chap. 15 (pp. 318–344). Boca Raton: Chapman and Hall/CRC.
79. Neuschmidt, O. (2012). Creating and Checking the TIMSS and PIRLS 2011 Databases. In M. O. Martin, & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
80. OECD (2003). *Education at a Glance 2003*. Organization for Economic Cooperation and Development.
81. Pallant, J. (2011). *SPSS survival manual: A step by step guide to data analysis using SPSS*, 4th edition. McGraw-Hill Education (UK).
82. Parker, P. D., Jerrim, J., Schoon, I., & Marsh, H. W. (2016). A multination study of socioeconomic inequality in expectations for progression to higher education: The role of between-school tracking and ability stratification. *American Educational Research Journal*, *53*, 6 –32.
83. Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*, 29–48.
84. Parker, P. D., Marsh, H. W., Guo, J., Anders, J., Shure, N., & Dicke, T. (2018). An Information Distortion Model of social class differences in math self-concept, intrinsic value and utility value. *Journal of Educational Psychology*, *110*, 445–463.
85. Pedhazur E. J. & Pedhazur Schmelkin L. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Psychology Press.
86. Kaur, J., Rana, J. S., & Kaur, R. (2009). Home environment and academic achievement as correlates of self-concept among adolescents. *Studies on Home and Community Science*, *3* (1), 13–17.
87. Last, J. M. (1988). *A dictionary of epidemiology* (2nd ed.). New York: Oxford University Press.
88. Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Erlbaum.
89. Rjosk C., Richter D., Hochweber J., Lüdtke O., Klieme E., Stanat P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: the mediating role of instructional quality. *Learn. Instr., 32*, pp. 63–72.
90. Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, *65*, 1–12.

91. Rutkowski D. & Rutkowski L. (2013). Measuring socioeconomic background in PISA: One size might not fit all. *Research in Comparative and International Education*, *8* (3), 259–278.
92. Rutkowski, L. & Rutkowski, D. (2010). Getting it 'better': the importance of improving background questionnaires in international large-scale assessment, *Journal of Curriculum Studies*, *42* (3), 411-430.
93. Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, *39*, 142–151.
94. Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25* (1), 54–67.
95. Salkind, N. (2010). *Encyclopedia of Research Design.* SAGE Publications, Inc., ebook.
96. Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407–441.
97. Sheldrake, R. (2016) Confidence as motivational expressions of interest, utility, and other influences: Exploring under-confidence and over-confidence in science students at secondary school. *International Journal of Educational Research*, *76*, pp. 50–65.
98. Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75* (3), 417–453.
99. Skolverket. (2009). *Vad påverkar resultaten i svensk grundskola? En kunskapsöversikt om betydelsen av olika faktorer* [What influences the outcomes of Swedish compulsory school? Review of the importance of different factors]. Stockholm: Skolverket.
100. Stankov, L., & Lee, J. (2017) Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence*, *61*, pp. 11–16.
101. Stankov, L., Lee, J., Luo, W., Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences, 22*, pp. 747–758.
102. Stankov, L., Morony, S., & Lee, Y. P. (2014). Confidence: the best non-cognitive predictor of academic achievement? *Educational Psychology, 34* (1), 9-28.
103. Steinmayr, R., Dinger, F. C., & Spinath, B. (2010). Parents' education and children's achievement: The role of personality. *European Journal of Personality*, *24*, 535–550.
104. Steinmayr, R., Dinger, F. C., & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality*, *26*, 335–349.
105. Teo, T. (2014). *Handbook of Quantitative Methods for Educational Research.* Rotterdam: Birkhäuser Boston.
106. UNDP (2014). *Human Development Report 2014. Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience*. New York: United Nations Development Programme.
107. Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review, *Educational Psychologist*, *39* (2), 111-133.

108. Volckmar, N. (2018). The Enduring Quest for Equity in Education: Comparing Norway and Australia, *Scandinavian Journal of Educational Research*, p.1-15.
109. Von Davier, M., Gonzalez, E., and Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series. Issues and Methodologies in Large–Scale Assessments, 2*: 9–36.
110. Vyas, S., & Choudhary, G. (2017). Relationship of socioeconomic status with frustration, self-concept, study habits and academic achievement of adolescents. *International Journal of Advanced Research and Development*, *2* (3), 46–51.
111. Weed, D. L. (1998). Commentary: Beyond black box epidemiology. *American Journal of Public Health, 88*, 12–14.
112. Wendt, H., Kasper, D., & Trendtel, M. (2017). Assuming measurement invariance of background indicators in international comparative educational achievement studies: a challenge for the interpretation of achievement differences. *Large-Scale Assessments in Education*, *5* (1), pp.1-34.
113. White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, *91*, 461–481.
114. Wigfield, A., & Eccles, J. S. (2000). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield, J. S. Eccles, & the Institute for Research on Women and Gender (Eds.), *Development of Achievement Motivation*, pp. 91–120. San Diego, CA: Academic Press.
115. Wu, M. (2005). The role of plausible values in large–scale surveys. *Studies in Educational Evaluation*, *31*: 114–28.
116. Wu, A. D., & Zumbo, B. D. (2007). Understanding and using mediators and moderators. *Social Indicators Research*, *87*: 367-392.
117. Wylie, R. C. (1989). *Measures of self-concept*. Lincoln: University of Nebraska Press.
118. Yamamoto, K., & Kulick, E. (2000). Scaling Methodology and Procedures for the TIMSS Mathematics and Science Scales. In M.O. Martin, K.D. Gregory, & S.E. Stemler (Eds.), *TIMSS 1999 technical report*. Chestnut Hill, MA: Boston College.
119. Yang Hansen, K. (2014). *Northern lights on TIMSS and PIRLS 2011: differences and similarities in the Nordic countries*. Köpenhamn: Nordic Council of Ministers.
120. Yang Hansen, K., & Gustafsson, J.-E. (2018). Identifying the key source of deteriorating educational equity in Sweden between 1998 and 2014. *International Journal of Educational Research*, https://doi.org/10.1016/j.ijer.2018.09.012.

Appendix 1. Home Educational Resources components (Martin et al., 2012).



Appendix 2. TIMSS 2011 scales for Students Confident in Science (self-concept), Students Value Science (utility), and Students Like Learning Science (interest).

Indicator[b]	Item	Factor
	Students Confident in Science scale separated into a variety of factors	
	I usually do well in <science>*	Self-concept
	<Science> is not one of my strengths[a]	Self-concept
	I learn things quickly in <science>	Self-concept
	I am good at working out difficult <science> problems	Self-concept
	<Science> is more difficult for me than for many of my classmates[a]	Peer-comparison
	<Science> is harder for me than any other subject[a]	Subject-comparison
	My teacher thinks I can do well in <science> <programs/ classes/lessons> with difficult materials	Teacher encouragement
	My teacher tells me I am good at <science>	Teacher encouragement
	<Science> makes me confused and nervous[a]	Anxiety
BSBGSCB BSBGSCC BSBGSCP	**Continuous indices of students confident in biology/chemistry/physics/earth science**	

BSBGSCE		
	Students Value Science (perceived utility of science)	
	I think learning <science> will help me in my daily life	Utility or extrinsic value
	I need <science> to learn other school subjects	
	I need to do well in <science> to get into the university of my choice	
	I need to do well in <science> to get the job I want	
	I would like a job that involves using <science>	
	It is important to do well in <science>	
BSBGSVB **BSBGSVC** **BSBGSVP** **BSBGSVE**	**Continuous indices of students' value of biology/chemistry/physics/earth science**	
	Students Like Learning Science (perceived interest of science)	
	I enjoy learning <science>	Interest (intrinsic value)
	I wish I did not have to study <science>[a]	
	<Science> is boring[a]	
	I learn many interesting things in <science>	
	I like <science>	
BSBGSLB **BSBGSLC** **BSBGSLP** **BSBGSLE**	**Indices of students' interest of biology/chemistry/physics/earth science**	

*<science> stands for biology/chemistry/physics/earth science, as in Sweden and Ukraine science is taught as separate subjects
[a] reverse coded
[b] – indicators are not given for all the items, because <science> is divided into four subjects for two scales given here, thus there are four indicators for some of the items, including four indices.

Appendix 3. Confidence Intervals (CIs) for Science Subdomains (Unstandardized)

3.1. Case of Sweden

```
CONFIDENCE INTERVALS OF MODEL RESULTS

                  Lower .5%  Lower 2.5%   Lower 5%    Estimate    Upper 5%  Upper 2.5%   Upper .5%

 TOTSCIE  BY
    PVCHE            1.000       1.000       1.000       1.000       1.000       1.000       1.000
    PVEAR            1.028       1.045       1.054       1.099       1.145       1.153       1.170
    PVBIO            0.977       0.989       0.995       1.027       1.059       1.066       1.078
    PVPHY            1.043       1.055       1.060       1.091       1.121       1.127       1.138

 Intercepts
    PVCHE            4.985       4.994       4.999       5.023       5.047       5.052       5.061
    PVEAR            5.160       5.169       5.173       5.197       5.221       5.225       5.234
    PVBIO            5.079       5.090       5.096       5.125       5.155       5.161       5.172
    PVPHY            4.920       4.934       4.941       4.980       5.018       5.025       5.040

 Variances
    TOTSCIE          0.522       0.539       0.547       0.592       0.636       0.645       0.662

 Residual Variances
    PVCHE            0.066       0.073       0.076       0.096       0.115       0.118       0.126
    PVEAR            0.097       0.102       0.105       0.119       0.133       0.136       0.141
    PVBIO            0.039       0.045       0.047       0.063       0.078       0.081       0.087
    PVPHY            0.072       0.079       0.083       0.102       0.122       0.126       0.133
```

3.2. Case of Ukraine

CONFIDENCE INTERVALS OF MODEL RESULTS

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
TOTSCIE BY							
PVCHE	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PVEAR	0.987	1.008	1.018	1.073	1.128	1.138	1.159
PVBIO	0.843	0.858	0.865	0.904	0.943	0.951	0.966
PVPHY	0.924	0.951	0.964	1.035	1.106	1.120	1.147
Intercepts							
PVCHE	5.003	5.030	5.043	5.115	5.187	5.201	5.228
PVEAR	4.841	4.866	4.879	4.947	5.016	5.029	5.054
PVBIO	4.830	4.852	4.864	4.924	4.984	4.995	5.018
PVPHY	4.908	4.937	4.951	5.027	5.103	5.118	5.146
Variances							
TOTSCIE	0.596	0.627	0.642	0.725	0.807	0.823	0.853
Residual Variances							
PVCHE	0.089	0.096	0.099	0.116	0.132	0.136	0.142
PVEAR	0.097	0.109	0.116	0.149	0.183	0.189	0.202
PVBIO	0.063	0.068	0.070	0.081	0.092	0.094	0.098
PVPHY	0.070	0.078	0.083	0.106	0.129	0.133	0.142

Appendix 4. Measurement Invariance for Latent Variables

4.1. Students Value Science continuous latent index.

Invariance Level		CFI	RMSEA	SRMR	Δ CFI*	Δ RMSEA
	Good fit	≥.95	≤ 0.08	≤ 0.06		
	Acceptable threshold	≥.90	≤ 0.10	≤ 0.10		
Configural		0.98	0.07	0.02	-	-
Metric		0.96	0.08	0.11	0.023	**0.009**
Scalar		0.95	0.07	0.13	**0.013**	**0.003**

*Model is invariant across two groups if absolute change in CFI (Δ CFI) ≤ .01, and absolute change in RMSEA (Δ RMSEA) ≤ .01. Values in bold – absolute invariance, italicized values – acceptable change.

4.2. Science Achievement continuous latent index (1st PVs of science subdomains used).

Invariance Level		CFI	RMSEA*	SRMR	Δ CFI	Δ RMSEA
	Good fit	≥.95	≤ 0.08	≤ 0.06		
	Acceptable threshold	≥.90	≤ 0.10	≤ 0.10		
Configural		1	0.05	0.003	-	-
Metric		1	0.08	0.08	**0.004**	0.03
Scalar		0.96	0.2	0.08	0.036	0.12

*Multiple imputation model type is applied to Science Achievement latent index used in the specified model, which reduces the measurement errors. Estimating measurement invariance together with TYPE=IMPUTATION is not possible in Mplus.

4.3. Parental Support categorical latent construct.

Invariance Level		CFI	RMSEA	SRMR	Δ CFI*	Δ RMSEA
	Good fit	≥.95	≤ 0.08	≤ 0.06		
	Acceptable threshold	≥.90	≤ 0.10	≤ 0.10		
Configural		0.90	0.16	0.05	-	-
Metric		0.89	0.12	0.07	**0.004**	**0.04**
Scalar		0.66	0.18	0.12	0.24	0.06

Appendix 5. Confidence Intervals for Structural Models (standardized)

5.1 Case of Sweden.

```
CONFIDENCE INTERVALS OF STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS


STDYX Standardization

                  Lower .5%  Lower 2.5%   Lower 5%    Estimate    Upper 5%  Upper 2.5%   Upper .5%

Effects from SES_SCL to TOTSCIE

  Total              0.384       0.394       0.400       0.428       0.456       0.462       0.472
  Total indirect     0.067       0.073       0.076       0.091       0.106       0.109       0.114

  Specific indirect

    TOTSCIE
    SCICONF
    SES_SCL          0.075       0.081       0.085       0.101       0.118       0.122       0.128

    TOTSCIE
    SCIVALUE
    SES_SCL         -0.020      -0.017      -0.016      -0.011      -0.005      -0.004      -0.002

  Direct
    TOTSCIE
    SES_SCL          0.292       0.303       0.308       0.337       0.366       0.372       0.382


Effects from GENDER to TOTSCIE

  Total             -0.009       0.000       0.005       0.028       0.052       0.057       0.066
  Total indirect     0.027       0.031       0.033       0.044       0.055       0.057       0.061
```

```
  Specific indirect

    TOTSCIE
    SCICONF
    GENDER             0.029        0.034        0.036        0.047        0.059        0.061        0.065

    TOTSCIE
    SCIVALUE
    GENDER            -0.008       -0.007       -0.006       -0.003        0.000        0.000        0.001

  Direct
    TOTSCIE
    GENDER            -0.050       -0.042       -0.038       -0.016        0.006        0.011        0.019


Effects from ETHNIC to TOTSCIE

  Total                0.099        0.109        0.114        0.141        0.168        0.174        0.184
  Total indirect      -0.001        0.003        0.005        0.015        0.026        0.028        0.032

  Specific indirect

    TOTSCIE
    SCICONF
    ETHNIC            -0.008       -0.004       -0.002        0.009        0.019        0.021        0.025

    TOTSCIE
    SCIVALUE
    ETHNIC             0.000        0.002        0.003        0.007        0.011        0.012        0.013

  Direct
    TOTSCIE
    ETHNIC             0.084        0.094        0.099        0.126        0.153        0.158        0.168


Effects from SUPPORT to TOTSCIE

  Total                0.083        0.094        0.100        0.130        0.160        0.166        0.177
  Total indirect      -0.065       -0.060       -0.057       -0.043       -0.029       -0.026       -0.021

  Specific indirect

    TOTSCIE
    SCICONF
    SUPPORT           -0.081       -0.075       -0.073       -0.058       -0.044       -0.041       -0.035

    TOTSCIE
    SCIVALUE
    SUPPORT            0.003        0.006        0.007        0.015        0.023        0.024        0.027

  Direct
    TOTSCIE
    SUPPORT            0.127        0.138        0.144        0.173        0.202        0.208        0.219
```

- **SUPPORT in Sweden’s model represents PSupport.*

5.2. Case of Ukraine.

```
CONFIDENCE INTERVALS OF STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS


STDYX Standardization

                  Lower .5%  Lower 2.5%   Lower 5%    Estimate    Upper 5%  Upper 2.5%   Upper .5%

Effects from SES_SCL to TOTSCIV

  Total             0.332       0.349       0.358       0.402       0.447       0.456       0.472
  Total indirect    0.114       0.123       0.128       0.154       0.179       0.184       0.194

  Specific indirect

    TOTSCIV
    SCICONF
    SES_SCL         0.099       0.108       0.113       0.138       0.162       0.167       0.176

    TOTSCIV
    SCIVALUE
    SES_SCL         0.002       0.005       0.007       0.016       0.025       0.027       0.030

  Direct
    TOTSCIV
    SES_SCL         0.177       0.194       0.203       0.249       0.294       0.303       0.320


Effects from GENDER to TOTSCIV

  Total            -0.004       0.010       0.017       0.055       0.092       0.100       0.114
  Total indirect   -0.009      -0.003       0.001       0.018       0.036       0.040       0.046


  Specific indirect

    TOTSCIV
    SCICONF
    GENDER         -0.018      -0.011      -0.007       0.013       0.033       0.037       0.044

    TOTSCIV
    SCIVALUE
    GENDER         -0.007      -0.004      -0.002       0.006       0.013       0.015       0.018

  Direct
    TOTSCIV
    GENDER         -0.024      -0.009      -0.002       0.036       0.075       0.082       0.096


Effects from ETHNIC to TOTSCIV

  Total            -0.196      -0.181      -0.173      -0.132      -0.092      -0.084      -0.069
  Total indirect   -0.047      -0.042      -0.039      -0.024      -0.010      -0.007      -0.002

  Specific indirect

    TOTSCIV
    SCICONF
    ETHNIC         -0.040      -0.035      -0.033      -0.019      -0.006      -0.004       0.001

    TOTSCIV
    SCIVALUE
    ETHNIC         -0.017      -0.014      -0.013      -0.005       0.003       0.004       0.007

  Direct
    TOTSCIV
    ETHNIC         -0.174      -0.159      -0.150      -0.108      -0.065      -0.057      -0.041
```

```
Effects from PSUPPORT to TOTSCIV

  Total            -0.019     0.002     0.012     0.067     0.122     0.132     0.152
  Total indirect   -0.112    -0.102    -0.096    -0.069    -0.041    -0.036    -0.026

  Specific indirect

    TOTSCIV
    SCICONF
    PSUPPORT       -0.171    -0.160    -0.155    -0.126    -0.098    -0.093    -0.082

    TOTSCIV
    SCIVALUE
    PSUPPORT        0.032     0.038     0.041     0.058     0.074     0.077     0.083

  Direct
    TOTSCIV
    PSUPPORT        0.047     0.068     0.079     0.136     0.192     0.203     0.224
```

Appendix 6. Mplus Inputs.

6.1. Input for testing ICC, two-level model.

TITLE: Intraclass correlation for models in Sweden / Ukraine;

DATA:
FILE IS TIMSS_2011_SWEDEN_SCIENCEPVSLIST.csv; OR !
UKRAINE_2011_SWEDEN_SCIENCEPVSLIST.csv !
TYPE IS IMPUTATION;

VARIABLE:
NAMES ARE NAMES ARE IDCNTRY IDBOOK IDSCHOOL IDCLASS IDSTUD
IDGRADE GENDER ETHNIC PSupport PSuppor2 PSuppor3 PSuppor4
IDPOP IDGRADER StudAGE IDSTRATE IDSTRATI TOTWGT HOUWGT
SENWGT SCHWGT CLASSWGT STUDWGT JKZONE
JKREP PVSCIE PVCHE PVEAR PVBIO PVPHY SES_scl
ValueBIO ValueCHE ValuePHY ValueEAR ConfBIO ConfCHE ConfPHY ConfEAR;

MISSING IS all (-99);
USEVARIABLES=PVCHE PVEAR PVBIO PVPHY SES_scl
ConfBIO ConfCHE ConfPHY ConfEAR ValueBIO ValueCHE
ValuePHY ValueEAR ETHNIC;
CLUSTER=IDSCHOOL OR ! IDCLASS; !
WEIGHT=HOUWGT;

DEFINE: PVCHE=PVCHE/100;
PVEAR=PVEAR/100;
PVBIO=PVBIO/100;
PVPHY=PVPHY/100;

ANALYSIS: TYPE=TWOLEVEL;

MODEL:

%WITHIN%
SciConf BY ConfBIO ConfCHE ConfPHY ConfEAR;

```
SciValue BY ValueBIO ValueCHE ValuePHY ValueEAR;
TotScie BY PVCHE PVEAR PVBIO PVPHY;

TotScie ON SciConf SciValue SES_scl ETHNIC;
SciConf ON SES_scl ETHNIC;
SciValue ON SES_scl ETHNIC;
SciConf WITH SciValue;
SES_scl ETHNIC GENDER WITH SES_scl ETHNIC;

%BETWEEN%

PVCHE PVEAR PVBIO PVPHY ON SES_scl;
ConfBIO ConfCHE ConfPHY ConfEAR ON SES_scl;
ValueBIO ValueCHE ValuePHY ValueEAR ON SES_scl;
ConfBIO ConfCHE ConfPHY ConfEAR WITH ValueBIO ValueCHE ValuePHY ValueEAR;

OUTPUT: STANDARDIZED;
```

6.2. Input of final structural model applied to Sweden's and Ukraine's sample.

```
TITLE: Mediator Model Sweden TIMSS 2011;   OR  Mediator Model Ukraine TIMSS 2011;

DATA, VARIABLES AND DEFINE COMMANDS SAME AS PREVIOUS INPUT 6.1.
CLUSTER = IDSCHOOL;

ANALYSIS: TYPE=COMPLEX;
       ESTIMATOR=MLR;          OR  ! BOOTSTRAP=5;
                                   REPSE=BOOTSTRAP; !

MODEL:
    SciConf BY ConfBIO ConfCHE ConfPHY ConfEAR;
    SciValue BY ValueBIO ValueCHE ValuePHY ValueEAR;
    TotScie BY PVCHE PVEAR PVBIO PVPHY;
    PSupport BY PSuppor1 PSuppor2 PSuppor3 PSuppor4;

    TotScie ON SciConf SciValue SES_scl GENDER ETHNIC PSupport;
    SciConf ON SES_scl GENDER ETHNIC PSupport;
    SciValue ON SES_scl GENDER ETHNIC PSupport;

    SciConf WITH SciValue;

    SES_scl GENDER ETHNIC PSupport WITH SES_scl GENDER ETHNIC PSupport;

    MODEL INDIRECT:
    TotScie IND SES_scl;
    TotScie IND GENDER;
    TotScie IND ETHNIC;
    TotScie IND PSupport;

OUTPUT: STdyx CINTERVAL sampstat;
```