THE GENDER GAP IN EDUCATIONAL ATTAINMENT IN THE FINNISH SCHOOL SYSTEM: A BROADER VIEW

Eero Syrjä
2377414
Master's Seminar and Thesis, 682436S
012.01.2022

Abstract

There is a large and growing gap between boys' and girls' educational attainment in Finland. This thesis examines the phenomenon in three different ways. The first section provides a historical overview of the state of the gender gap in Finland using statistical data collected by Statistics Finland and shows how the gap is growing each year. The second section performs a systematic review of research literature on recent, peer-reviewed studies of Finnish students. The third section analyses a recent questionnaire by the Finnish National Agency for Education regarding the progress of a pilot programme for L2 learning. The results show that there is an almost universal advantage to girls, ranging from temperament, language skills, eating habits, developmental speeds to PISA performance. These differences were observed from a very young age, and the differences increased with age. The thesis further examines the current attitudes of Finnish teachers seen in the responses of a questionnaire performed by the Finnish National Agency for Education, showing a large lack of concern or awareness of the effects of gender on educational performance. The combination of the gender gap as it exists now, the lack of any large scale directed research, and the apparent blindness of the effects of gender by the respondents to the questionnaire all show a large, systemic sexist effect in favour of girls in the Finnish educational system.

Tyttöjen ja poikien koulumenestyksessä on suuri ja kasvava ero Suomessa. Tämä opinnäytetyö tutkii kolmea eri puolta tästä ilmiöstä. Työn ensimmäinen osa tarjoaa historiallisen katsauksen ja tilastollisen kuvan tämänhetkisestä sukupuolijakaumasta Suomessa. Työn toisessa osassa suoritetaan systemaattisen kirjallisuuskatsauksen uusista, vertaisarvioiduista tutkimuksesta suomalaisista oppilaista. Työn kolmas osa analysoi uutta opetushallituksen suorittamaa kyselyä osana L2 oppimisen kärkihanketta. Tulokset osoittavat että tytöt suoriutuvat paremmin melkein kaikissa tilanteissa, luonteesta kielitaitoon, syömistottumuksiin, kehitysnopeuteen ja PISA-kokeisiin. Nämä erot havaittiin hyvin nuorista alkaen ja ero sukupuolten välillä kasvoi iän myötä. Lisäksi tämä opinnäyte tutkii suomalaisten opettajien tämänhetkisiä asenteita sukupuoleen kasvatuskontekstissa. Kyselyssä monella ilmenee vähättelevä tai suoraan hylkäävä asenne sukupuolen vaikutukseen koulumenestyksessä. Tämänhetkinen tilastollinen tilanne, laaja-alaisen tutkimuksen puute ilmiöstä sekä kyselyyn vastaavien näennäinen asennoituminen kaikki muodostavat kuvan suuresta, systemaattisesta seksistisestä vaikutuksesta tyttöjen eduksi suomalaisessa koulujärjestelmässä.

Table of Contents

Contents

Introduction

Education is seen as of great importance in Finland, requiring high standards for teachers, with teaching seen as one of the more respected professions in the country, in addition to requiring a master's degree for teacher's qualifications. This emphasis on education was closely tied to the success of post-war Finland, as the country historically had very little industry and wealth, especially when compared to its neighbours. It was the concerted and deliberate efforts to educate the populace that allowed Finland to develop quickly, even as it struggled with the costs of the Second World War, being one of the few nations to pay its war reparations in full, all without the support of the Marshall plan.

Today, Finland enjoys a high standard of living, and is one of the best places in the world to be a woman (Best countries for Women, Papadopoulos), regularly placing in the top 10 in different multinational surveys, and the World Economic Forum ranking Finland as third in its 2020 Global Gender Gap Report (World Economic Forum 9). There is, however, a problem in educational trends in Finland that show a significant and persistent difference in educational attainment between in favour of girls.

When considering an entire country's educational system as a whole, clearly there are many different factors that influence the outcomes it produces. These factors have been the focus of studies for decades, with one large subset of study being the impact of gender in educational attainment. This imbalance on educational outcomes has been seen in practically every European country, as PISA 2018 saw an average gap of 30 score points—with standard deviation being 100 points—in reading in OECD countries, and the gap existed in "all countries and economies that participated" (OECD, "Finland – Country Note" 4, 6). This topic has seen increased focus as of late, as the imbalance between the genders in Finland is growing by the year. Finding a comprehensive explanation as to why the situation is developing has proven difficult, and previous efforts have focused on many different possible factors such as the students themselves, the environment, schools, families, or the curriculum. There is a lack of a systematic overview of all the factors that influence educational outcome, and this thesis aims to provide some insight to the overall picture.

This thesis examines the current educational attainment gap that exists in Finland. The aim is to provide an overview of the current state of gender imbalance in educational attainment, to review the existing research on the topic, and to identify the key contributory factors.
There are three questions in relation to those topics that this thesis aims to answer:

1. How has the educational gap between the genders developed in Finland over the last 50 years?
2. What does recent, peer reviewed research indicate are the possible causes for the gender imbalance?
3. What are Finnish teachers' attitudes towards the question of gender and its influence in L2 education?

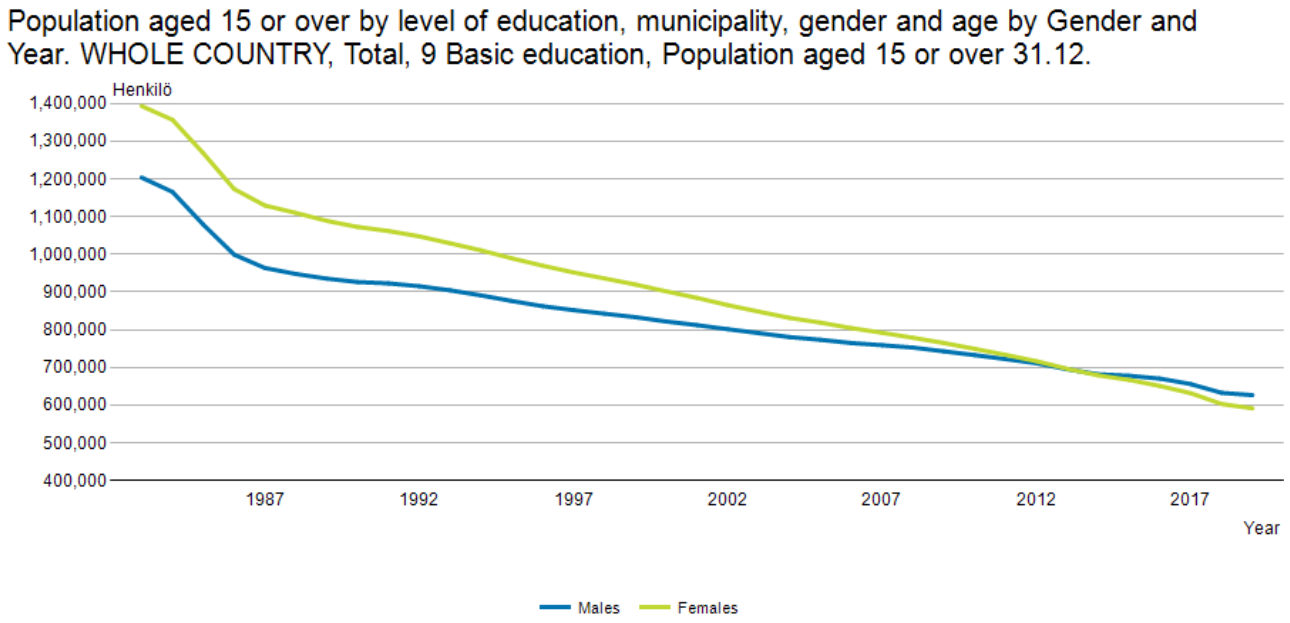
Current statistics concerning the growth of the educational gender gap in Finland are examined in the first section. By providing information regarding the scale of the issue currently in Finland, along with historical trends, the first section clarifies the urgent need to address the problem as the disparity is worsening. The second section of the thesis is a systematic review of literature to find existing studies that can provide contributory factors for this large gender gap in educational attainment. A systematic review allows for a thorough, and importantly, repeatable overview of current literature as it relates to differences between the genders in school attainment in Finland, and from there a qualitative analysis is performed on the results. This paper does not include a meta-analysis, as although typical when performing a systematic review, the varied nature of the review articles makes such an analysis impossible. A review protocol will be presented, followed by an overview of the results of the search, and an analysis of the information extracted from the studies that were selected. The third section is an analysis of a recent interim report questionnaire from the Department of Education of Finland, focusing on an analysis of an open-ended question regarding teachers' perceptions around the effect of gender in L2 language learning. An overview of the function and nature of questionnaires as a data gathering method is presented, with a thorough analysis of the answers to allow for contextualisation. Finally, a discussion draws connections between the three sections and their findings, giving an overview of possible explanations as to why girls are doing so much better in school than boys.

As this paper is trying to find explanations regarding gender effects in the education system, those students who fall outside of this binary definition of gender will be excluded from consideration here. This is unfortunately due to time restrictions, and availability of public data. This is in no way whatsoever an attempt to minimize their importance, or to ignore the difficulties they face in trying to progress through the educational system, but rather an acknowledgement of the inability of this thesis to properly account for those of special circumstances. The acceptance of those who do not fall into the stereotypical gender division is low in society at large, with Finland still to this day requiring sterilization for those who wish to undergo gender-reassignment surgery. The experience

of those who go through mandatory education while suffering from gender dysmorphia deserves its own full considerations.

Section One - The educational gender gap in Finland: An overview

The gender imbalance in school attainment in Finland has been growing larger every year, with equivalence in absolute numbers between the genders momentarily reached around the new millennium, with some variance as to which year depending on the level of educational attainment (Figures 1-3). This follows a longer-term trend of girls and women outpacing boys' and men's performances in education that began to appear in the 1980's. The effect of this imbalance can be seen in the statistics on the educational attainment of the population that are shown in the following graphs. These graphs are exclusionary: they only show the number of people that are exclusively at a given level of educational attainment. This is clear in Figure 1, with the overall decline in the number of people who stop their education at 9 years shows that a larger portion of the total population go on to higher levels of secondary or tertiary education:



Source: Educational structure of population, Statistics Finland

Figure 1

This decline in absolute numbers is significant, as the population of the country increased from 4.5 million in 1970, to 5.5 million in 2019. Unfortunately, Statistics Finland only has data available

online going back to 1970, which is the range for the graphs shown here. The data shown is, however, enough to show the changes in performance of the genders in the educational system.

In Figure 1 there are a few interesting historical points, namely the large, and seemingly equal, decline of those who stop at basic education from 1970 to 1985, and then the faster decline of females from there on. Starting in 1970, there were 200 000 more women than men who remained at the level of basic education, and it took some 42 years for equivalence to be reached. Since then, women have outperformed men. Finishing grade 9 education in Finland is effectively universal: it is mandated by law, and the number of students who did not complete their mandatory education varied between 0.17% and 0.75% between 1999 and 2018 (OSF). We can thus make the inference that a having a smaller number of women who stop at this level of education means they went on to a higher level of education, a claim that will be shown with the following graphs and data.

This effect of gender in school performance can be seen in larger, international studies as well. In the latest PISA national report on Finland, the results of PISA 2018 are discussed, and in it, the extent of this gap between 15-year-old students in Finland was revealed. On average the gap between the genders was 52 score points in reading, and when compared to PISA 2009, boys' performance fell (OECD, "Finland – Country Note" 6). This means that the gap between the genders in Finland was 60% larger than the OECD average, which was already at 30 score points in favour of girls.

This gap is also evident in both mathematics and sciences: Finland was one of the few countries participating in PISA testing in which girls outperformed boys in mathematics, although here the gap is at its smallest, at only 6 score points. Even then, the advantage that girls in Finland had over boys is larger than the average advantage of boys in the OECD countries, which is at 5 points. (OECD, "Executive Summary" 16; OECD, "Finland – Country Note" 4).

The gap between the genders widened in the sciences section, with girls outperforming boys by 24 score points in Finland, when the average in OECD is only 2 score points. In understanding what these values mean, it should be noted that PISA does not use minimum or maximum grades for their scoring, instead using normalised scales, with the overall OECD average going to 500 points, and a standard deviation being 100 points (OECD, "PISA FAQ").

This means that, on average, Finnish girls were half a standard deviation better than boys at reading, and a quarter better in the sciences. This average, however, does not tell the whole story. The

distribution of scores is an important factor, as it influences our understanding of the differences between the genders as well as finding possible explanations for this effect. In the 2018 PISA Results Volume II, chapter 7, this specific distribution is expanded upon. There, the report stated:

> ...in all countries/economies, the first decile of the performance distribution amongst boys was significantly lower than that amongst girls. In Finland, Israel, Jordan, Malta, Norway, Qatar, Saudi Arabia and the United Arab Emirates, the 10% lowest-performing boys scored at least 60 points lower than the 10% lowest-performing girls.

This effect of the lowest scoring boys' being significantly worse than the lowest scoring girls' is a very important one, as the aforementioned report points out that Finland was one of the small group of countries where the mean level of reading for boys was above the OECD average.

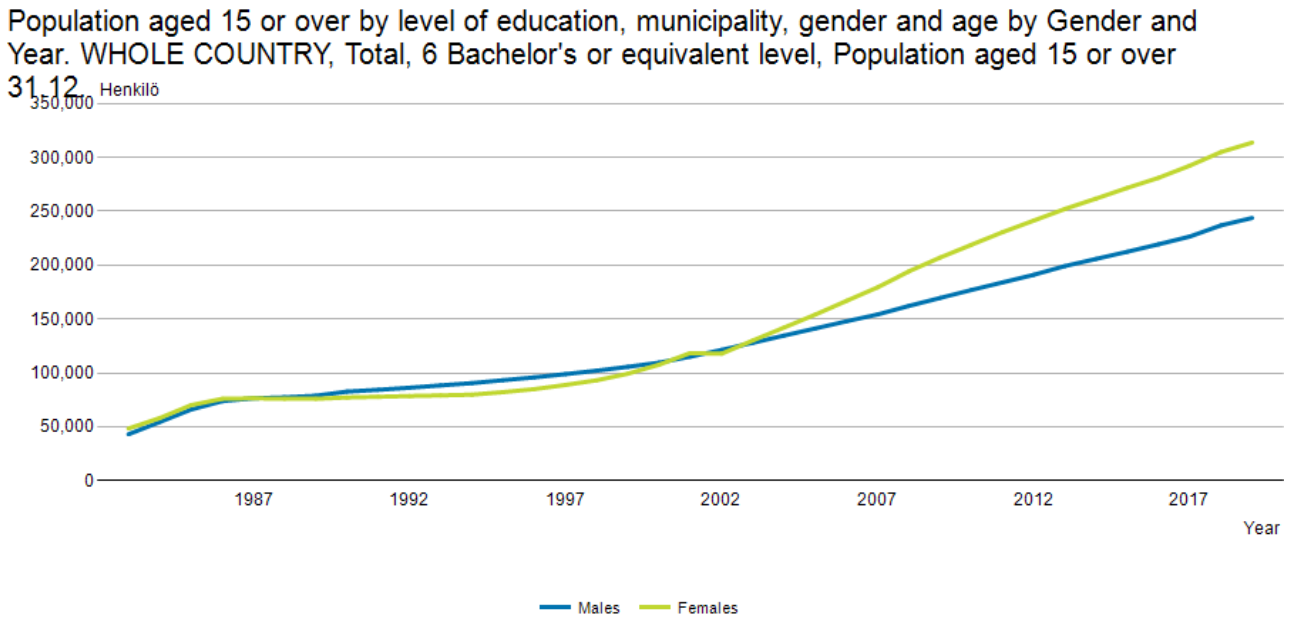
Finland has one of the largest gaps in reading, a significant gap in sciences, and an advantage in mathematics for girls' performance in PISA 2018. The fact is that 15-year-old girls significantly outperformed boys, but this level of a gap is not seen in every country. There were countries in PISA 2018 that had reduced the gap between the genders when compared to earlier PISA scores, as was noted in the OECD report on gender imbalance of the results from 2018:

> In 17 of those countries/economies (of the 36 participating countries that showed improvements in reducing the gender gap), the narrowing of the gender gap in reading performance was due to an improvement in the performance of boys. In five of those countries/ economies, namely Estonia, Ireland, Macao (China), Peru and Singapore, boys and girls in 2018 scored higher in reading than their counterparts did in 2009, even as the gender gap between them shrank during the period.

This is evidence that there is no biological or fundamental reason for the difference between the genders' performance, but rather, it is an artifact of the education system (OECD, "Girls' and Boys' Performance in PISA" 146).

There is only one conclusion: There is a systemic, sexist effect in the Finnish education system that favours women over men. In Finland, secondary and tertiary education access is based on what appear to be meritocratic systems: entrance exams are point based, and transparent. Grades from earlier schooling affect a portion of those chosen for a given school program, but a portion of available spots are based solely on the entrance exam, to enable those who did not do well in school an opportunity to get into higher education. The reality is, however, different. The statistics regarding gender imbalance in higher education in Finland are stark, and while the systems around admittance into higher education appear to be meritocratic and gender neutral, the outcome clearly shows that this is not the case.

As an example, the effect can be seen at the lower levels of tertiary education, with the BA level educational makeup of the population being as follows:

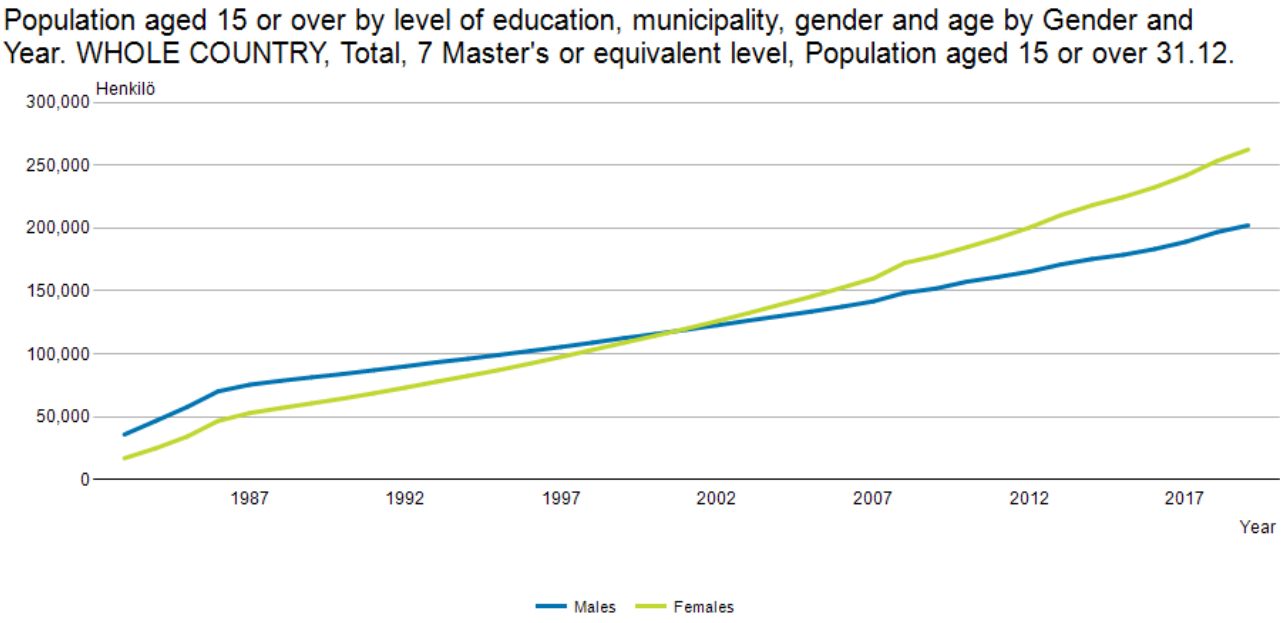


Source: Educational structure of population, Statistics Finland

Figure 2

Here we see the point of equivalence at roughly the year 2000. The period from 1970 to 2000 is more closely matched than with the other categories, but after 2000, the difference appears and is currently widening rapidly.

In upper tertiary education the ratio of males to females grew from effectively equal numbers in 2000, to 30% more females than males in just 19 years:



Source: Educational structure of population, Statistics Finland

Figure 3

This graph shows a similar development pattern from figure 1 where the growth rate is closer to equal from 1970 to 1985, after which there is a marked difference between the genders. When compared to basic education, the point if equity comes earlier (2000 vs. 2013). Effectively all university degrees in Finland fall under MA level of education, which this table shows. Clearly, there is a marked difference in the performance of genders in attaining this level of education, and the gap in 2019 was a bit over 60 000, a larger number than the sum of both genders with a MA level of education when the statistics started in 1970, which was 52 952.

If we analyse what these three different tables show us in total, we can see that in the current system women significantly outperform men in terms of access to, and competition in secondary and tertiary education. The effect is clear and continuous, and the rate of separation between the two genders is increasing with each year.

As has also been shown, the scale of the gap is large and widespread in most developed countries. While it would be extremely difficult, if not impossible, to try and find a definitive causal explanation for a phenomenon as multifaceted as this, this paper will attempt to find as many possible explanations to the question at hand.

Section Two – A Systematic Review of the Literature

2.1 Methodology

This paper adapts the PRISMA system of systematic reviews to find possible explanations to the large and persistently increasing gender gap in educational attainment. While this method is primarily meant for the review of literature in medicine, here it is adapted to fit this review of literature regarding language and education.

A systematic review is a protocol of literary review that is meant to be a repeatable by others to reach the same results. The reason for using this specific method is that it allows for a thorough search for explanatory and descriptive research on the topic.

This paper will build on and expand on the review that was done in *Girls rule, boys drool! A Systematic Review of Literature on the Effects of Gender on Educational Outcomes* (Syrjä, 2016). There, a review of two databases for the years of 2000 to 2016 from ProQuest's ERIC database and EBSCO's all databases was performed. This thesis will search those two databases and look for new results in the four years since the previous search was performed. Thus, the focus of this systematic review will be to find all relevant peer-reviewed studies that have been published in the years of 2017-2020 from the two databases mentioned earlier.

Additionally, this thesis will include two new databases in the search: Scopus and Fennica. The inclusion of these two databases is to increase the chance of finding Finnish studies, as the focus of this paper is specifically on research related to Finland. In these two database searches, the years included will be from 2000 to 2020, to give equivalence to the searches performed in the earlier work.

The aim of the systematic review is to find all studies that offer explanations regarding the gender gap in educational attainments. The studies will be limited to those regarding Finnish students, as the question is complex enough in a single societal and national context. To attempt to find commonalities in an international setting would be outside the scope of this thesis.

The results of the search were checked for duplicates. After this, a reading of the abstract, and possibly some of the body of the text was done to find those studies that are relevant to the research questions in this thesis. These relevant articles were read closely, taking note of the specific aspects

of the study such as methodology, participants, and results, as well as summarising the relevance of said study to this thesis' research questions. An analysis of common themes and similar causal links between articles was done, grouping the studies based on three categories and finally a summary of evidence found was done and discussed.

The search itself will use Boolean operators to specify and limit the scope of results in the search. These operators are logical operators and allow for a very exact search of results.

The search terms used are as follows:

1. (gender* OR sex*)
2. AND (gap OR differen*)
3. AND (education* OR schoo*)
4. AND (finlan* OR finnis*)
5. AND (languag* OR readin*)
6. NOT (disab* OR dropou*)

The results of this search will be limited to peer reviewed, English language studies that are available in full text online, within the dates depending on which database is being searched.

The first search term is used to find studies that relate to gender, as this is directly relevant to the research question of this thesis, and any study that focuses on the difference between the two genders will be relevant to it. The second search term has two logical operators that require one of either gap or difference terms to be found in the results. This is to enable the inclusion of any studies that have anything to do with gender differences or gaps. The third search term limits the results to those that have to do with schools or education, as gender studies are a broad field that, without this limitation, certainly would give many more results. The fourth search term limits the results to those that include Finland in some manner. This does not exclude international studies, or those that use representative groups from other countries, but does limit it to those that have at least some Finnish participants.

This limitation does exclude by necessity some relevant studies, for example Lisbeth and Glenn in "Outliers: Upper secondary school students who read better in L2 than in L1" found that of those students who scored higher in English than their mother tongue, two thirds were boys. This impact of English as an alternative language and its effect on school performance was not among any of the results of the systematic review, and since the aforementioned study was performed in Norway, it is

excluded in this thesis. Although the inclusion of such apparently closely related studies might be warranted, the differences between educational systems of Finland and Norway makes its inclusion here unfounded.

The fifth search term focuses the results to language and reading. This is justified since the largest gap in performance has been shown to be in reading skills. The sixth search term removes those results that have to do with stopping schooling, or those of special needs. While there exists a gender gap in dropping out of school, and in the way that special needs present across the genders, it is a specialty field that requires its own consideration, and as such, is outside the scope of this thesis.

2.2 Study selection

Studies will be selected based on their relevancy to the main question of this paper, that is, the gender gap in education. As the search itself is not limiting the methodologies or target populations of these studies, the selection will be done with a careful reading of the abstract, and where necessary, the article itself. The selected studies vary greatly in their methodologies and target populations, as was seen in Syrjä 2016, and as such, this systematic review will do a qualitative analysis rather than a quantitative one.

2.2.1 Data collection process

The collection of articles was done using the Boolean search phrase shown previously. This search was performed on the four databases mentioned, and was limited to fully available, peer reviewed sources from 1.1.2000-28.2.2021. The results of these searches were exported to RefWorks citation manager, where duplicates were removed. After this, the sources were reviewed individually with the relevant ones included, and those that were not relevant to the research question excluded.

2.2.2 Data items

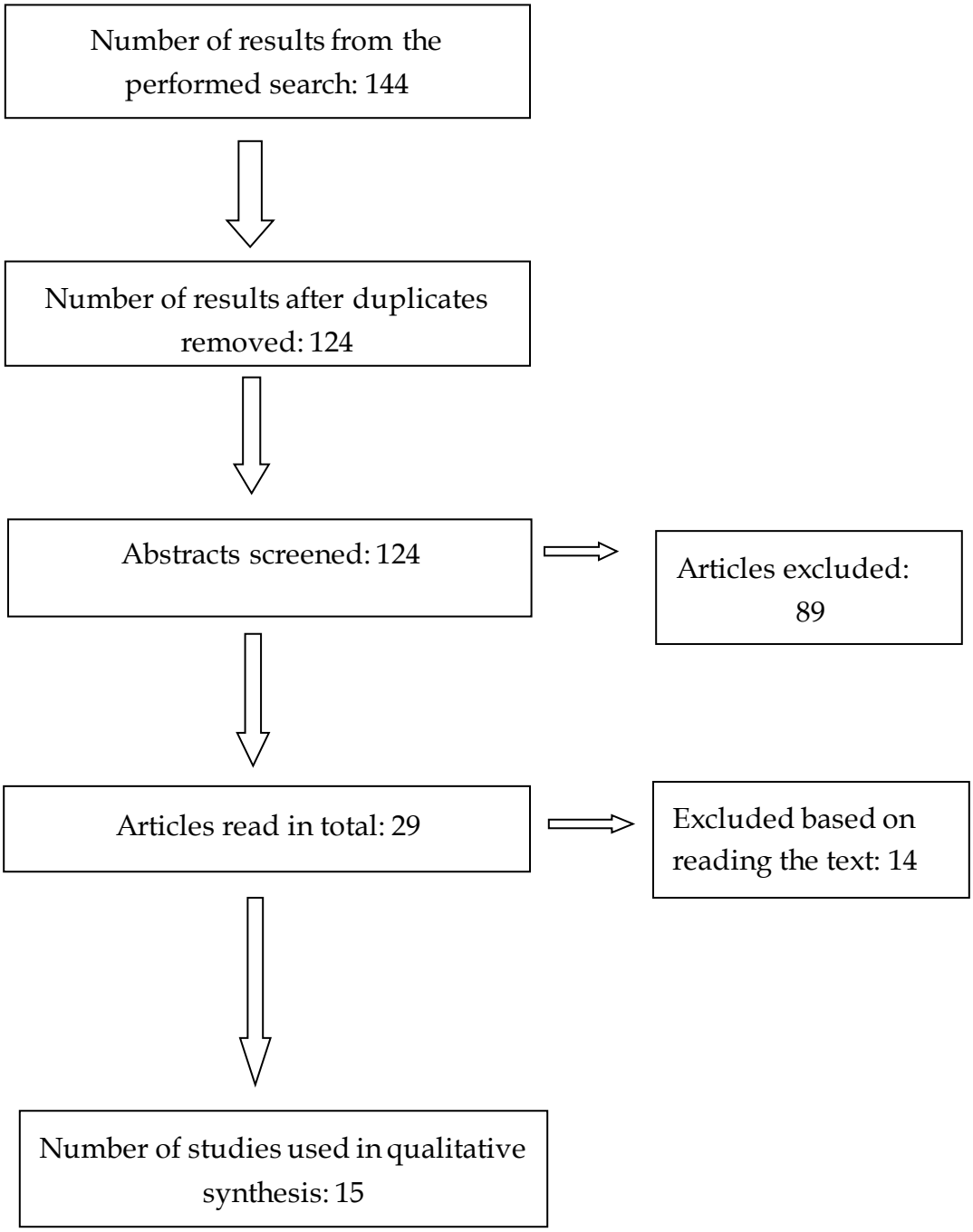
This study is aiming to find causes, explanations, and compounding factors with regards to the gender imbalance found in the Finnish school systems. As such, the data collected from the selected studies will focus on study subjects, methodology, and results. Since only those studies that have to

do with the gender gap in Finnish school system will be included, the aim is to have data about Finnish students and their relative performances.

2.2.3 Risk of bias

The main risk of bias for this paper is the limitation of possible results in the search of studies. The Boolean phrase was developed to find the most relevant results possible. This restrictive structure of the search phrases does exclude most results. This is by design, as the number of studies that have the term education or gender is so large as to be impossible to try and analyse in a study of this kind. A fundamental limitation of research of this kind is the inability to change the variable being studied. There is no ethical way to change a student's gender to see if their school performance would increase, and as such, the results of these studies are always inferential and statistical, with a lack of casual direction. This is a fundamental limitation of studying this problem, and cannot be avoided, but must be considered when considering the results and analysis of the studies.

2.3 Study Selection



2.3.1 Selected studies

AUTHORS:	METHOD:	GENDER EFFECT:	STUDY DESIGN:	PARTICIPANTS:	RESULTS:
MÄKI, H. S., ET AL.	Multiple due to different types of skill being measured.	Advantage to girls' performance, predictive value unable to explain gap.	Longitudinal	154 first graders at the start of first year.	Preschool skills predicted writing skill development
DE OLIVEIRA FIGUEIREDO, REJANE AUGUSTA, ET AL.	Self-reported through questionnaire.	Girls eat healthier meals than boys	Cross-sectional	10569 under 18's in Finland.	Girls make up a higher portion of healthy eaters, and boys of unhealthy eaters.
FRÖJD, S. A., ET AL.	Anonymous questionnaire.	Girls more likely to self-report depression.	Cross-sectional	2516 7th to 9th grade students.	Self-reported depression and difficulties in school predicted self-reported school performance, but direction of cause was unclear.
KORPILAHTI, P., A. KALJONEN, AND E. JANSSON-VERKASALO.	Demographic data collection, questionnaires, infant language development measurement in labs	Gender was found to be a major factor in language development, with an advantage to girls	Longitudinal	The larger cohort of 9811 mothers and 9936 children, and a follow-up group consisting of 1797 mothers and 1827 children	10 risk factors for poor skills in language comprehension at 36 months were found, they include father's employment status, socioeconomic factors, and gender

AUTHORS:	METHOD:	GENDER EFFECT:	STUDY DESIGN:	PARTICIPANTS:	RESULTS:
KRKOVIC, K., ET AL.	Questionnaire using Likert scale	Girls were rated better than boys' by teachers of both genders	Cross-sectional	1979 6th graders, 188 teachers	Teachers' evaluation of students' chance of school success and skill in language were influenced by the gender of the student, outside objective performance.
LEPOLA, J.	Individual tests administered by the researchers	Girls' task orientation was higher at every point of examination	Longitudinal	157 pre-schoolers	Boys with low comprehension skills were shown to develop more socially dependant behaviour in grade 1,
MANU, M., ET AL.	Comparison of kindergarten skills to PISA performance at grade 9	Kindergarten skills predicted more variance for boys than girls in PISA	Longitudinal	1839 kindergarteners and 1010 grade 9 students	Kindergarten skills predicted PISA performance
MULLOLA, S., ET AL.	Analysis of three studies	Girls are seen as much more ideal students	-	3212 students and 221 teachers	Girls were shown to have more desirable temperamental traits than boys
OPPERMANN, E., ET AL.	Questionnaire using Likert scale	Girls were less likely to be in maths motivated profile	Longitudinal	292 grade 2 students, 345 grade 3 students	Majority of students belonged to highly motivated group, with

					variation between genders based on motivational groups.

AUTHORS:	METHOD:	GENDER EFFECT:	STUDY DESIGN:	PARTICIPANTS:	RESULTS:
ORELLANA, P., ET AL.	Self-reported	Girls' motivation to read in Finland higher at every level than boys' motivation.	Cross-sectional	K-2 grade students, 302 from Chile, 701 from USA, 307 from Greece and 254 from Finland	Reading motivation increased each year for all students, with girls starting with a higher level, and maintaining it.
PEURA, P.I., ET AL.	Likert-scale questionnaire	Boys reported higher self-efficacy in digital reading	Cross-sectional	1327 students from grades 2 to 5	Positive correlation between self-efficacy beliefs and reading fluency
PESU, LAURA, ET AL.	Questionnaire to both students and their parents	Boys reported higher values in mathematics, girls in literature	Longitudinal	231 students at grade 7, 221 of them at grade 9, as well as 221 mothers and 191 fathers	Self-concept of ability for mathematics and literature declined from grade 7 to 9. Parents belief at grade 7 predicted grade 9 self-conception
PÖYSÄ, S., ET AL.	Questionnaire	Boys showed worse engagement in four categories, girls were never worse.	Cross-sectional	307 grade 7 students	Four distinct subgroups were found based on engagement profiles, with high engagement correlating with a high GPA, but no causal direction was shown
SILINSKAS, G., ET AL.	In class test of students, and a questionnaire to parents	No gender difference in the effect of maternal teaching	Longitudinal	1460 mother-child dyads	Positive association between maternal controlling behaviour, emotional values,

TORPPA, M., ET AL.	In-class testing and a questionnaire	Significant gender differences in favour of girls	Cross-sectional	1309 9th graders	and high involvement levels with regards to reading skills Significant difference between genders in PISA-style reading testing, as well as leisure time reading habits

2.3.2 Discussion of extracted data

The studies selected for this paper were decidedly heterogenous in terms of their research focus, methods, and subjects. For clarity, they have been divided into three groups: The home environment, the school environment, and studies regarding the students themselves. This division is particularly relevant to discussions around the education system as outside actors have a limited ability to influence the inner workings of families, and as such, the practical value of studying the school environment itself is higher in discussions regarding the educational outcomes of students. The studies will be discussed, possible over-arching themes and connected results will be presented and the most important findings along with a connecting synthesis will be done to show the relevance of the findings to the research question of this paper.

2.3.3 Studies regarding pre-school and home environments:

The studies included in this section are the most varied in their research targets, ranging from maternal reading practice to the development of communication skills in infants, and the eating habits of children. When looking at the overall influence on educational success, the home environment is significant, but changes to it are difficult to achieve through policy. There is a strong societal and governmental support system in place for new parents in Finland, with most day-care centres and preschools being run by the local municipalities, along with parental education and maternal support. When considering the influence of gender in the family situation, there are gender norms and roles that are passed from generation to generation, with traditional ideas of masculinity and femininity passed on from generation to generation. These attitudes and roles certainly do play a role in the educational outcomes of students, and the following studies will show results regarding this. Unfortunately, influencing the home environment with policy or governmental action is challenging, as parents have a fundamental right to raise their children as they see fit, outside of abuse cases. Thus, the importance of studying the school environment along with teachers' role in it is evident, to find possible explanations for, and solutions to the educational attainment gap.

The effects of home environment on school performance are of import to the research questions of this paper. De Oliveira Figueiredo et al. in "Identifying Eating Habits in Finnish Children: A Cross-Sectional Study" studied the way in which young people in Finland eat, using a rather large sample size of 10 569 respondents. The study divided the respondents into three categories: unhealthy

eaters, fruit and vegetable avoiders, and healthy eaters (de Oliveira Figueiredo, et al. 3). With regards to gender differences, boys made up a larger portion of unhealthy eaters, and girls of healthy eaters in the study, but despite the better overall eating habits, girls had a higher incidence of irregular breakfasts than boys (de Oliveira Figueiredo, et al. 4). The study does note, however, that the pattern of unhealthy eating in boys has been shown in multiple earlier studies, and that girls are more likely to eat healthy foods than boys (de Oliveira Figueiredo, et al. 8). The effects of good nutrition on school performance have been studied elsewhere, but the causal link between school performance and the gendered effect of nutrition found in this study is not shown here. An inference between the two can be argued, but as there is no link shown in the study, the value of such an analysis is limited. Of more importance to this paper is the pattern of findings surrounding difference in the habits of boys and girls outside the school, as this is a recurring theme in the studies analysed here.

In "Identification of Biological and Environmental Risk Factors for Language Delay: The Let's Talk STEPS Study", Korpilahti and Jansson-Verkasalo examined how newborns and infants' language skills developed from birth to 3-years-old. They used a large cohort of parents and children, with a focus on finding risk factors to language development, including environmental effects like parents' socioeconomic backgrounds, employment status etc. Gender was found to be a major factor, with a significant statistical difference in developmental speed using the Reynell Developmental Language Scales III (Korpilahti and Jansson-Verkasalo 31). This is not to say that gender is the only risk factor that influences the child's language development, as father's employment status at the time, as well both parents' social class were also identified as risk factors. However, the study found that:

> Gender was the only biological factor that predicted poor skills in language comprehension at 36 months (Korpilahti and Jansson-Verkasalo 32)

Further, one of the larger predictors of language acquisition is the gender of the child, with girls having statistically significant better development at 24 months, and this difference continues to the age of 36 months (Korpilahti and Jansson-Verkasalo 28). That a difference between the genders is already seen with children of such young age, and specifically a difference that is not influenced by outside factors, but is a biological one, is of significant import to the research questions of this paper. Considering the large sample size of this study, the control of other variables and the length of observation over three years, the findings do indicate that a fundamental difference in development exists. Earlier studies such as Lim et al. in "Preferential Detachment During Human Brain Development: Age- and Sex-Specific Structural Connectivity in Diffusion Tensor Imaging (DTI) Data" have shown that there are differences in the development of the brain between the

genders, but the question is how the school system should take this difference into account. If boys develop at a different rate, should gendered schools be considered? Do they need extra support, and how large of an effect do the studies show? The following studies show additional data regarding the development of pre-school children's skills.

This development of reading skills and specifically the role of the mother in it were studied in Silinskas et al.'s "Maternal Teaching of Reading and Children's Reading Skills in Grade 1: Patterns and Predictors of Positive and Negative Associations". Here, the researchers tested first graders at the start of and at the end of the first year of school, as well as sending questionnaires to the mothers regarding maternal teaching of reading.

The study found that the lower a child's reading skills at the start of grade 1, the higher the mother's involvement (Silinskas et al. 59). This increased maternal teaching negatively predicted their child's skills at the end of grade 1 (Silinskas et al. 59). The four groups found were:

> (1) High skills and no contribution of teaching (12%), (2) Low skills and positive contribution of teaching (14%), (3) Low skills and no contribution of teaching (22%), and (4) Average skills and negative contribution of teaching (52%) (Silinskas et al. 59).

These groups are with regards to the predictive value of maternal teaching in the child's reading skills at the end of the first grade, so negative contribution means that the teaching was unable to improve the child's reading (Silinskas et al. 60). The parent's emotional response to homework situations was also shown to be a factor in reading skill development, with positive attitudes being important with an increased effort in maternal teaching (Silinskas et al. 63). This effect was hypothesized to be due to parental emotional signalling about the nature of homework to the child, where a positive approach gave the child an impression of overcoming difficulties with effort, and negative emotions showing children that homework "should be avoided" (Silinskas et al. 63). Most importantly, the study showed no gender differences between any of the four subgroups, thus showing that boys and girls both benefit equally from maternal teaching efforts (Silinskas et al. 64).

Since no gender differences between the subgroups of child-mother dyads were shown, it means that parental involvement is of equivalent value to low performing boys and girls. The positive association between maternal controlling behaviour, emotional values, and high involvement levels with regards to reading skills shows that those students who are lower performing can be supported to better their skill levels, even when controlling for pre-grade 1 skills. The question, then, is why other studies find such stark differences between the genders reading performance. If the ability of

both boys and girls to improve their skills is equivalent, what is the effect that causes such a large difference in outcomes?

Differences in how students of different genders do homework and read in their leisure time is another important factor in home environments, and this was studied in Torppa et al.'s "Why do Boys and Girls Perform Differently on PISA Reading in Finland? the Effects of Reading Fluency, Achievement Behaviour, Leisure Reading and Homework Activity". The study measured both objective performances using PISA testing material, and the self-reported differences in homework and reading behaviours at home. The study found that reading fluency, PISA Reading performance, and homework activity all showed "significant gender differences...in favour of girls" (Torppa et al. 131). There were also differences in what kind of material was read, with boys preferring comics, and girls reading more books and magazines (Torppa et al. 131).

Earlier studies have also shown that the gender typical approaches to school success are that girls succeed through effort, but boys succeed due to giftedness (reference here from BA). Torppa et al. show here that, at least as far as reading scores in PISA go, reading fluency is the largest influencing factor, with smaller effects from leisure reading, and time spent doing homework. The issue, as the researchers rightly point out, is that this does not provide any reason or explanation for this gap, but rather states that the reason why boys do worse in PISA reading testing is because they are worse readers. The size of this gap was significant, with 2.4 times higher risk in PISA reading scores and 4.4 times higher risk in reading fluency for boys to belong to the "lowest tail" of scores (Torppa et al. 131). The fact that this gap exists is the issue, as PISA scores are a symptom, rather than the cause itself, as they reflect the difference in genuine ability that is gendered. This seems to follow the other studies covered in this paper, that show a gap in reading ability starting from kindergarten, that seems to continue all the way through the school system.

The influence of early- and pre-school skills in future performance is a subject that was covered multiple times in the studies reviewed in this paper. Perhaps the most thorough and the one with the longest study period was Manu et al. in "Kindergarten Pre-Reading Skills Predict Grade 9 Reading Comprehension (PISA Reading) but Fail to Explain Gender Difference", as it used data from a 10-year longitudinal study. The 10-year study first measured kindergarteners' skills in "phonological awareness, rapid naming, letter knowledge, vocabulary and listening comprehension" in 2006, and this data was then used as a possible predictor of how a portion of those students would perform in PISA tests (Manu et al. 759). This predictive value would be further examined by whether there was a difference in the predictive value of boys and girls, if there was a difference between the genders

already at kindergarten level, and if the differences seen in PISA can be fully explained by these differences (Manu et al. 757-758).

Overall, the study measured the predictive effects of kindergarten skills as explaining "27% of the variance in reading comprehension...among boys, and 18% among girls" found in PISA scores (Manu et al. 764). Regarding the gender ratio of those who fall in the lower percentile of readers in this PISA sample, the results were quite stark, with the ratio of boys to girls being "1.99:1 for percentile 30, 2.83:1 for percentile 15, 8.75:1 for percentile 5, and 7.67:1 for percentile 3" (Manu et al. 765). The study also found that while there is some variance in the kindergarten level skills in pre-reading between the genders, this variance is not one sided, and the size of the effect is small (Manu et al. 765).

Perhaps most damningly, the study found that only 27% of boys' and 18% of girls' variance in PISA results were explained by kindergarten skills, indicating a difference in the way that boys and girls develop and learn through the compulsory education system. While boys' reading in PISA scores is, on average, above the level of the OECD average, the fact that girls are doing so much better in these tests indicates a systemic issue with the Finnish school system. Furthermore, considering that boys have a larger predictive factor between their pre-reading skills and their PISA scores would seem to indicate that girls have a better chance of improving their proficiencies during grade school, or that boys are more "punished" for their lack of prereading skills from kindergarten. This is additionally worrisome as other studies show an advantage to girls' skills in preschool level language skills. If the predictive value of preschool skills for boys' PISA performance is higher than girls, and their preschool skill levels are worse, it means that in addition to having a harder time improving their skills through their school careers, they start with a disadvantage. The gender ratios of bad readers are also alarming, as the ratio of being in the bottom 5% of the results was almost 9:1 of boys to girls. This result supports other studies that have shown a difference in the distribution of students' performance that is not exemplified by simply looking at average values.

In summary, the studies here show a varied field of different effects that have gendered differences. Girls develop their language skills as infants faster, they have healthier eating habits, they score significantly better when tested using the PISA test's reading sections, and their leisure reading habits were different from boys. The size of the effect varied, but when there was a difference, the direction was universal: in favour of girls. These findings will be put in the context of schools in the following sections.

2.3.4 Studies regarding students

Studies that focus on the students themselves will be summarized here. A fundamental limitation regarding research in this field is the inability to change the variable being studied, namely gender. It would clearly be unethical to see if a students' school achievements were improved by changing their gender, and as such, all studies in this category are looking at differences in how boys and girls act, learn, behave, study, and develop in schools using correlational data as though it indicates cause-effect relationship. The function of the following studies is to enable a clearer understanding of the gendered differences in students to allow for better planning of schools and education, to enable students to flourish in their school careers regardless of their gender identity.

Mäki et al. in "Predicting Writing Skill Development with Word Recognition and Preschool Readiness Skills." studied the effects of preschool skills on the development of students' reading and writing skills up to the end of grade 3. The study found that girls had "better mechanics of writing and they wrote more coherent stories than boys" (Mäki et al. 665). This difference between the genders was not fully explained by their predictive model, that attempted to have the results at grade 3 explained by the preschool skill levels. Even when considering the different skill levels of the genders when moving from pre-school to first grade—which themselves are open to the question of why this difference exists—the study was unable to explain why there was an increased difference between the genders in grades 2 and 3 (Mäki et al. 666). Furthermore, the authors note that they could already see a difference between boys' and girls' skills when in pre-school, which might indicate that the difference in writing skills is starting before kindergarten (Mäki et al 667).

2.3.5 Studies Regarding Motivation

The differences between pre-kindergarten skills were also discussed in Lepola's "The Role of Gender and Reading Competence in the Development of Motivational Orientations from Kindergarten to Grade 1", which looked at the effect of pre-reading skills in children's motivational orientation in grade 1. The study used both pre-school testing and grade 1 testing, in combination with a teachers' evaluation of the students' motivation to find associations between these two factors. The study found that at every point of time examined, girls' task orientation was graded as higher than boys', as well as in socially dependent behaviour (Lepola 225), but the latter effect disappeared in Grade one. When looking at covariables between language comprehension, gender

and time, the study found that boys' who had "low comprehension skill" would progressive show more socially dependant behaviour in grade 1, whereas boys' with normal or high comprehension skills and girls' of all skill levels showed either a consistent level of, or a reduction in the same behaviour from preschool to grade 1 (Lepola 230).

Lepola's study also found that a child's levels of pre-reading skills in phonetics or language comprehension correlated to their motivational orientation in grade 1, and that a good competency in either acts as a protective measure on their developing of what were rated as negative behaviour patterns in socially dependant behaviour and motivational orientation (Lepola 233-234). The study also found that while girls were rated as better than boys at every level, there was no gender difference in the development of either task orientation or socially dependant behaviour from preschool to grade 1 (Lepola 233-234). Lepola shows here that both prereading skill levels and gender do impact the development of the students' task orientation, but gender did not show a difference in socially dependant behaviour.

Furthermore, it was noted that boys with low comprehension levels in language and phonetics were at risk of developing worse patterns for both behaviours from preschool to Grade one, where all other groups showed either consistent levels, or improvements in their behaviour. As Lepola notes, however, there are some possible factors that influence the conclusions that can be drawn from this, namely, the differences between the teachers' relationship and students of different genders, as well as how the difference in the ways in which girls seek help when compared to boys might colour the grading of their behaviour by the teachers (Lepola 235-236). This, combined with the fact that the study does not provide any specific information as to the gender makeup of the three skill groups that were assigned at the start of the study, makes it somewhat difficult to draw many inferences. We can see that girls are rated better than boys by their teachers, and that boys specifically showed a larger effect in worsening behaviour if their prereading competencies were low to begin with, but we are given no information as to whether there was a difference in the initial state of skills based on gender.

The question of why boys react in a different, and in an apparently worse, way to the experience of preschool and Grade one teaching when compared to similar, low competency level girls, is of significant import as well, as presumably both groups are participating in the same teaching, going through the same materials, with similar peers. One possible explanation is that there are gendered behavioural norms that influence the perception of teachers in grading their students' actions.

The subject of motivation was further covered in Orellana et al.'s "Motivation to Read in Grades K–2: A Cross-Cultural Perspective". Here, the researchers looked at the development of motivation to read over a school year, from kindergarten to grade 2. The researchers found that there is an increase in motivation to read from year to year, and between the start and the end of the school year (Orellana et al. 434). In addition, the study found that there were significant differences between the genders' motivation, with girls rating their motivation to read as higher at every level of study, with the smallest gap in kindergarten, and largest at grade 2 (Orellana et al. 431). This shows that the development of motivation through the first two years of school is heavily favouring girls, and that the difference increases over time. These results might explain why Torppa et al. showed a difference in the reading habits of boys and girls, as a better developed language skill, as well as a higher level of motivation towards reading at an early school age might reflect even as late as grade 9.

Motivational factors in early primary school education were further studied by Oppermann et al. in "Elementary School Students' Motivational Profiles Across Finnish Language, Mathematics and Science: Longitudinal Trajectories, Gender Differences and STEM Aspirations". In this study, the students were asked to "rate their intrinsic values in maths, Finnish and science" with a Likert-scale like scale, with regards to their emotional attitude towards the subject, their belief in their own ability to successfully perform studies related to it, as well as the students' "dream job", with the answer being categorized into groupings to see possible factors between students' ratings and their desire to have a STEM career (Oppermann et al. 5). The students' ratings were divided into three separate categories: low motivation, high motivation, and high maths motivation (Oppermann et al. 7). There was a rather sizable difference between the categories, with the sizes being 78 for low motivation, 181 for high motivation, and 34 for high maths motivation (Oppermann et al. 7).

The self-perception of a students' abilities is important to the way they will approach tasks in school, as well as their performance in the subject, as these factors clearly influence motivational and the students' attitudes regarding tasks. The findings here regarding self-reported levels of ability and skill are somewhat limited due to the young age of the participants but are still illuminating regarding the differences between the genders' views of both themselves, and the subjects in question. Oppermann et al. acknowledged the limited nature of the data they worked it, as they studied students in grades 2 and 3, the findings related to the gender makeup of the motivation groups is still relevant as it shows that there is a gendered difference in mathematics motivation already this early on, and how gender roles influenced goals regarding dream jobs. Regarding gender differences, girls were less likely to be in the maths motivated profile, and that

there was no difference between low and high motivated profiles' gender balance (Oppermann et al. 7). The probability of remaining in a specific group differed, with girls being more likely to stay in the high motivation group than boys, and with boys more likely to stay in the maths motivated group, or to move to it from the other two, than girls were (Oppermann et al. 7). Although the study did find that girls were less likely to move to, or to belong to the maths motivated group, this group had a lower level of motivation in Finnish than the low motivation group; it is characterized by the high motivation in maths and low motivation in other subjects (Oppermann et al. 10). Thus, the fact that girls are in higher numbers in the high motivation group is an advantage, as the reported levels and attitudes towards all three subjects were highest in this group (Oppermann et al. 10).

The fact that these effects are shown at this young of age is worrisome, as it would seem to indicate that a larger portion of boys are falling out of the high motivation group that enables their belief in freely achieving any possible future career that they want. The high motivation group was by far the largest group out of the three, being 160% bigger than the other two groups combined, but unfortunately the study does not show what the different group makeups were based on gender. The study does show, however, that girls are more likely to belong to the high motivation profile, and to stay in it, and the fact that this effect is already seen from grades 2 to 3 is certainly relevant to the research questions of this paper.

2.3.6 Studies Regarding Self-efficacy

The role of self-perceived skill level in a task on motivation regarding said task was examined in Lepola's study, and Peura et al. add with their study to that discussion by showing a clear correlation between self-reported efficacy and objectively measured literacy skills in "Specificity of Reading Self-Efficacy among Primary School Children." In the study, they examined the self-perceived efficacy levels students from grades 2 to 5 using a questionnaire. This questionnaire used a 7-point scale ranging from "I'm totally certain I can't" to "I'm totally certain I can" for the given questions (Peura et al 503). Additionally, the children were administered a reading fluency test using three time-limited tests (Peura et al 504). The study did not show a large difference between the genders, noting that the differences between the genders were "small to moderate" and stating that more study was required for making any claims (Peura et al 511).

The study found positive correlation between self-efficacy beliefs and measured reading fluency (Peura et al. 510). It further showed that efficacy beliefs became more nuanced with increased age,

and the ability of students to identify their own skills in a more granular way developed along with their skills (Peura et al. 510). Unfortunately, the study looked at gender specific differences only in passing and did not analyse the differences in any meaningful way, only noting that more research was needed to make any claims. The study also found a correlation between self-efficacy beliefs and reading fluency, specifically noting that: "the more children believed in their capabilities in reading activities related to daily life, the better their reading fluency skills were" (Peura et al. 510). Additionally, it found that reading motivation is tied to reading fluency, which in turn is tied to "independent reading practice" (Peura et al. 510-511), arguing that self-efficacy beliefs influence the motivational factors of students' approach to practicing reading, and through that, influence the fluency level and motivation level of reading (Peura et al. 511).

The association between self-perceived skills, different specificity levels of skills and literacy skills is of import to this paper's research question and showing the existence of this effect as early as grade 2 has implications for how we understand the development of both the students' self-image as well as their motivation and actual abilities with regards to reading fluency.

Other studies in this paper have discussed the different ways in which girls and boys read, and the materials that they read. The reported association between self-efficacy, reading fluency and "independent reading practice" is very significant when considered with the differences in how, and what, the genders read.

Research around this topic was further discussed in Pesu et al.'s "The Development of Adolescents' Self-Concept of Ability through Grades 7-9 and the Role of Parental Beliefs". Here, the researchers looked at the development of self-concept through middle-school, as well as the parents' view regarding their children's abilities. The study found that self-concept declined both in maths and literature from grade 7 to 9, and while girls had a higher level of self-concept than boys, the rate of change was similar for both genders (Pesu et al. 99). This demonstrates the lack of gender difference in the development of self-concept in mathematics and literature from grade 7 to 9, but the lack of difference in the development does not mean that there is no difference in the level of reported self-concept, and here the results follow what could be called expected values, where boys reported higher values in mathematics, and girls reported higher values in literature (Pesu et al. 99). Unfortunately, there was no data on the relative performances of the genders at any measurement point, as the focus of the study was on the changes of self-concept and its relation to parental beliefs. Consequently, there is no direct comparative data to show any correlation between self-concept, gender, and performance; but there is a lack of gendered difference in the development of

these self-concepts, from which we can infer that both genders in this study were influenced evenly in these subjects.

These findings seem to show a contradictory result when compared to researched covered earlier, that showed a difference in the development of motivation and task-orientation in elementary school students. This could indicate a difference between elementary and middle schools and their effects on the students, but regardless of the possible reasons, the study shows again a higher value of self-concept for girls. The fact that the changes in these values over the years was similar to both genders seem to follow along with the study regarding mother's involvement in learning to read, where a similar equivalent effect was seen.

2.3.7 Studies Regarding Student Engagement

Differences in how girls and boys engage in school activities is the focus of Pöysä et al.'s study "Adolescents' Engagement Profiles and their Association with Academic Performance and Situational Engagement". The study measured both overall and situational engagement, with overall engagement using three categories of engagement: behavioural, emotional, and cognitive (Pöysä et al. 4). The study found "significant gender differences" in four areas: behavioural engagement, reading comprehension, reading fluency and situational engagement showed higher values for girls (Pöysä et al. 5). The study further found that there are four distinct subgroups formed by their level of overall engagement: high overall engagement, high future goals and mid overall engagement, mid overall engagement, and low overall engagement levels (Pöysä et al. 9), with the mid overall engagement group the largest. Furthermore, the prevalence of boys in the low overall engagement group was higher than girls, and conversely boys were less represented in the high and high and future goals groups (Pöysä et al. 9). There was also an association between engagement groups and GPA, with higher levels of engagement manifesting a higher level of academic performance (Pöysä et al. 9).

The role of engagement and its influence on how a student interacts with and relates to situations in school is a very important one. Higher overall engagement having an association with higher GPA is an important factor, however, the direction of causation is not shown here, and cannot be inferred by the data in the study: are students who are motivated better at school, or are students who are better at school more motivated? Regarding the research question of this thesis, the study provides more data showing the prevalence of girls in higher motivation groups and boys in low motivation

groups, but it does not give any direct values, nor does it make any gendered analysis regarding engagement differences outside of the given value differences between boys and girls in four categories of behavioural engagement, reading comprehension and fluency and situational engagement (Pöysä et al. 5). There were a total of 15 different categories that were measured, and only in four were there differences between the genders, and these value differences were not put into the context of different motivation groups, so the inferences that can be drawn from this data are limited. Regardless, it was shown that boys are more likely to have lower overall engagement towards school, perform worse in reading skills, and have a higher level of disaffection in school situations. This supports the other studies reviewed in this paper, as girls' performance in literature is a universally common theme in the academic research included in this paper.

Another important factor studied was the association between school performance and motivation in school. Fröid et al. in "Depression and School Performance in Middle Adolescent Boys and Girls" studied the association between self-reported depression and school performance, comparing GPA and depressive values. The study has found that self-reported depression is a predictor of academic performance, but the causal link is not always clear: a student might be doing badly in school because of depression, or they might be depressed because they are doing badly in school (Fröid et al. 493).

The study further found that there were differences between the genders, as there was a significantly higher incidence of girls reporting depressive symptoms than boys. This result is a significant departure from most results found in this paper, as it showed some advantage to boys. However, this apparent advantage to boys is not necessarily clear-cut, as while the study itself shows that school performance is linked to depression, other studies covered in this paper—as well as PISA scores and other sources cited— show how girls are outperforming boys in schools at every level. Therefore, if school performance and depression are positively associated, as the study shows, there should be more depressed boys than girls, as boys do worse in school than girls. The fact that the opposite result has been shown would seem to indicate differences in how the genders answer questions regarding their mental states, but the study does not consider this aspect, and as such, the question is left unanswered.

2.3.8 Studies Regarding the School Environment

The areas of the studies reviewed in this section are those that look at factors that are external to the students, with the two studies here focusing on teachers and their relationship with, and influence on, the students. When speaking of gender and its influence on school outcomes, the effect of teachers as graders and educators is clearly significant. All mandatory school grades in Finland have, as a part of their grading system, the behaviour of the pupil as a factor: those who behave well and actively participate get a higher grade than those who behave badly and are not active. The inclusion of this behavioural grading reflects the societal expectations and attitudes towards what a student is, and what an "exemplary" student behaves like. Considering this normative feedback of behaviour in classrooms, the role of teachers in affecting the gendered differences in outcomes must be clarified and examined.

Krkovic et al. approach this question in their study "Teacher Evaluation of Student Ability: What Roles do Teacher Gender, Student Gender, and their Interaction Play?" by comparing the results of a questionnaire given to teachers, and their students objective academic performance. The students were given tests, and the results of these tests were used to objectively measure performance, and this data was then compared to the teachers' evaluations to find interactions between the students' and teachers' genders. Teachers are influenced by gender stereotypes as well, as they are humans and a part of the society that created these stereotypes. The study found that teachers of both genders evaluated girls' performance in language and school success higher than boys' performance. Both male and female teachers did this, and this effect was seen even after objective performance was controlled for, but it was missing in mathematics evaluations (Krkovic et al. 252-253). The study further showed that while teachers did not have a cross-gender effect where they would have a bias towards their own gender, there was a general bias in favour of girls by teachers of both genders (Krkovic et al. 254). Importantly, this effect is seen even after controlling for objective performance, meaning that the results show a general preference for, and an overestimation of, girls' school performance. This "ideal" student type was also shown in Mullola et al., that seems to indicate a pervasive preference for a specific type of learner, and a specific manner of behaviour, that is outside any kind of objective performance or academic ability.

In Mullola et al.'s "Same Temperament, Different Appreciation: Temperament's Influence on School Achievement Depends on Student's and Teacher's Gender and Teacher's Age.", three studies were analysed with regards to the relationship between temperament, student teachability, teacher and student gender, and the interaction between these factors. (Mullola et al. 87). The students'

temperament was measured, and this was compared to school achievement using the students' school report grades (Mullola et al. 88). The three analysed studies show a clear difference between the perceived temperaments of students based on their gender, with boys' being seen as having higher levels of activity, negative emotions, lower inhibitions, and distractibility; where girls are seen as having higher persistence, better mood, better educational competency, cognitive ability, motivation, and maturity (Mullola et al. 89).

These significant differences can influence the teachers' perception of the students' teachability and general preference towards a type of temperament, and this is especially significant as a part of a teachers' evaluation of a student is their perceived attitude towards schoolwork, and this perception influences the student's grade. The study also found that the teachers' gender influenced how they perceived students of the same gender, with both showing a preference to their own gender (Mullola et al. 89). Furthermore, male teachers specifically perceived their male students as having "better" scores in temperament but were more critical of female students when compared to female teachers' evaluations (Mullola et al. 90). Additionally, there was a noticeable effect of male teachers' age and their perception of male students' temperament, with a higher rating of inhibition, and a lower rating of mood for male students (Mullola et al. 90), but similar effects were not seen for female students of older male teachers, nor from female teachers for students of either gender (Mullola et al. 90).

Mullola et al. points to three specific factors that affect the teachability that a teacher views a student with: task orientation, personal-social flexibility, and reactivity (81-82). The chapter goes on to note that in all three of these categories, studies have shown that the teachers' perception is that boys do worse than girls in them. This, however, is not an indicator of the capabilities of students, but rather a reflection of the perception that teachers' have regarding the teachability of their students, due to the way they act in school (Mullola et al. 81-83). This effect is further confounded by the fact that similar behaviour can be seen in a very different way depending on the gender of the student, and the gender of the teacher, as well as if they are of the same gender, or of different genders (Mullola et al. 83,85).

Strikingly, the study was unable to determine the effect of teachers' gender on perceptions of students' abilities in Finnish language courses, as there were no male teachers of Finnish in the data (Mullola et al. 94). This is not entirely unexpected, as in a 2019 publication, the Finnish National Agency for Education showed that the gender balance of teachers in mandatory education in

Finland, with 70-80% of teachers being female, with the only exception being school principals, where the gender balance was almost even (OPH 11).

When comparing the results from Krkovic et al. and Mullola et al., Krkovic et al. found that there was no bias difference with teachers of both genders preferring girls over boys, but in Mullola et al. the difference is in perceived temperaments specifically, and the study still showed a significant difference in the way that girls' behaviour was perceived as better. The discussion around teacher perceptions and preconceptions regarding teachability, academic performance and behaviour is significant, as teachers have an enormous influence on both the students' self-perceived ability and in their motivation, as well as objective feedback through grading.

2.4 Conclusions

This section has included data and analysis of 15 studies selected from over 120 results of a systematic review on this topic. There was a single study that showed clear apparent advantage to boys, namely Fröid et al., but this departure from the norm is most likely influenced by the self-reported nature of the study, and the differences in gender roles influencing how depression is talked about. Mental health in men in general is lacking, with attitudes towards using mental health care as generally negative and almost taboo, and it would seem likely that this is the reason for this anomaly. If we were to assume the study is valid on its own merits, it would mean that we must conclude that even though girls are more depressed than boys, and that depression is associated with school success and GPA, girls are still significantly outperforming boys in school. This would indicate that the systemic sexist effect of schools would be so significant as to overcome the advantage boys have in mental health, but it is difficult to say which interpretation of the situation is correct, and more study is needed here.

The results of the rest of the studies speak for themselves. There were studies from pre-natal development all the way to PISA scores at 15 years old, and at every step, the advantage was to girls. To be clear, this advantage was not in every category, and was not always significant, but the singular direction of the effect is striking: nowhere were boys ahead of girls, not even in what would be traditionally seen as male dominated fields such as mathematics or sciences.

Additionally, the impact of teacher attitudes towards students and the perception of behaviour was a significant point of concern, as both Krkovic et al. and Mullola et al. showed a preference for girls' behaviour beyond objective performance. The role of teachers in motivating students and enabling learning is well established, and if teachers are prejudicial towards the way girls behave as a subjective preference, that will be reflected in their daily guidance and feedback that they give to their students.

As has been shown in this section, the number of recent, peer reviewed studies about Finnish students is surprisingly low, with only 15 studies applicable out of the 124 results that the systematic search produced. In terms of answering the question set out for this section, there certainly were many different effects and situations studied, with numerous different age groups, sample sizes and research methods. The lack of practically any advantage to boys' performance is of significant concern, as when looking at worldwide situation such as with the PISA 2018 gender performance, the gender gap closed in half the participating countries, and in roughly half of those it was due to both genders improving their performance.

Large systems such as the educational systems are difficult to change, and as has been shown by the number of resulting studies in this paper, the factors that influence such a system are varied. It would be very difficult to attempt to give a singular causal explanation for a phenomenon such as this, even if more time and proper research was done on the topic, as the number of variables to account for is simply too large. This does not, however, mean that there is nothing to be gained. By studying the impact that specific aspects of the school system have on students, we are able to better the system itself, and guide the way in which we participate in it.

Specifically, the role of teachers and their preconceptions regarding the differences between girls and boys in school is a topic that should be researched further, as it has the largest impact with the lowest cost. Teachers are rightly the focus of much attention and expectations in the Finnish school system, and as such, if there are unrealised biases that are colouring the way they teach their students, this should be examined and clarified.

Section Three: A Case Study of Teachers' Perceptions

As shown in the previous sections, there is ample scientific research that shows the influence that a student's gender has in educational outcomes. The effect was seen across all studies discussed in the systematic review, not a single study showed an advantage in favour of boys' performances, with an equivalence between the genders being the best outcome. This difference is seen from a very young age (For examples see Mäki et al; Korpilahti, Kaljonen and Jansson-Verkasalo; and Lepola), and the effects are evident throughout the entire educational system. The systematic review covered all relevant results from four major academic databases covering 20 years' worth of research. Four categories were created to create groupings that showed the relevant results from the chosen studies. Out of all the studies selected for the review, only two were relating to teachers' attitudes and perceptions of student performance, Krkovic et al. and Mullola et al. The results from these two studies were contradictory, and as they were a small minority of the studies covered, there is a clear need for researching teachers' perceptions and attitudes surrounding the role of a student's gender in the school environment.

This section of this thesis examines a questionnaire regarding a pilot initiative run by the Finnish National Agency for Education testing the effects of early childhood L2 learning. The question under analysis specifically relates to how the teachers participating in the programme perceive gender differences in their teaching. This determination has great practical merit to the discussion around the impact of gender in educational outcomes, since if teachers are predisposed towards a specific approach, understanding or gender, or ignorant of the impact of students' gender in this context, it would increase the gap in educational outcomes that we see.

3.1 Teachers' perceptions of gender differences in L2 learning

As a part of the Finnish government's government programme (hallitusohjelma) under Prime Minister Juha Sipilä in 2015, the Finnish National Agency for Education began an initiative (kärkihanke) in 2017, which aimed to change the starting point of early language education in Finnish schools. The purpose of the initiative was to study the effects of lowering the starting age of L2 and L3 learning in schools, as well as to study what effects increased early language learning might have on the students' language learning throughout their school careers.

The initiative was continued in 2018 with two groups of participants: Group A which continued pre-existing initiatives from 2017, and Group B which began participating in the scheme. The aims for the initiative in 2018 were stated as:

1. To bring early language education into earlier grades, with the aim of taking advantage of early childhood language learning abilities.
2. To fund initiatives that aim to develop students' motivation regarding language studies, with a focus on long-term studying and seeing it as a part of lifelong development.
3. To use funding to encourage study environments that recognize the role of different languages in the setting of early and elementary education.
4. To increase awareness around language learning, and to use up to date, research-based information regarding the advantages of early language learning, as well as to critically analyse current structures around how students are presented information regarding language studies (Opetushallitus Hakutiedote).

81 schools were selected to continue the initiative in Group A, and the Finnish National Agency for Education allocated €3 427 000 between the participants. The participants were required to prepare an interim report on the progress of their initiatives by January 2019 for the Finnish National Agency for Education, and all 81 participating schools did so. The specific nature and execution of the individual initiatives were left to the participants, which included many different types of institutions, including municipalities and universities.

This interim report included a questionnaire, which included the question (12):

> Have you noticed a difference between the genders in language study and learning? What differences have you noticed? (Opetushallitus Väliselvitys)

The answers to this question were provided in open format, with no personally identifiable information. Since the question is open-ended, there is significant variation between the answers, with large variance in answer length and content. Some of the answers were in Swedish, but most were in Finnish.

3.2 Questionnaires: a common tool for collecting research data

Questionnaires are a method of data collection, typically used in quantitative research as the method lends itself to easy collection of data from large populations. As with any other form of research, questionnaires require consideration of the ethics involved with the initiative, as the participants are not "passive data providers" (Cohen et. al. 377). This active participation by the subjects is of central importance when designing a questionnaire, as the resulting data is dependent on how the respondents react with and understand the questionnaire.

This aspect of questionnaires illuminates a fundamental property of what kind of information a questionnaire gathers: it measures the participants subjective understanding of the question, or what they are willing to answer regarding it. The form that questionnaires take can vary greatly and is mostly determined by the number of participants in the study, with a higher number of participants making a more structured and closed design a better choice (Cohen et al. 381). In such a case, much care must be given to the design of the questionnaire, typically with a pilot run and thorough planning to make certain that the full range of possible answers are included in the questionnaire itself (Cohen et al. 382). The other side of this structured/unstructured scale is an open questionnaire that is cohesive and gives direction in the data collection but allows for "respondents to reply in their own terms" (Cohen et al. 382). There are differences in how these two different styles of questionnaires are analysed, as they give the researcher different kinds of data to analyse due to their structural differences.

One of the more common types of questionnaires use rating scales questions with the Likert scale, which can ask the participants to rate their agreement of a sentence with a scale of values, typically between four and seven (Cohen et al. 388). The form of the questionnaire will impact the kind of analysis that can be done on the resulting data, with the answers given to the question at hand diverging significantly from the kinds of analysis done on a Likert-scale questionnaire. Open-ended questions allow for a wide range of answers, some of which might not expected by the researcher, and as such the resulting data must be coded, organised, and then analysed in a different way when compared to a typical close-ended questionnaire.

Making questionnaires even more challenging to properly design is the issue of variables. The function of a questionnaire is to measure opinions, perceptions, and beliefs of the respondents. Questionnaires can also be used as tool in study design to measure changes brought about by

adjusting dependent variables, but the questionnaire being examined here is measuring the perceptions of the respondents to the given question.

The questionnaire analysed in this section of this thesis was designed and executed by professionals at the Finnish National Agency for Education. It consisted of 15 questions, of which 11 were open ended questions. Additionally, the questionnaire included a section on the current progress of the pilot programmes, with a dedicated section for reporting the specific forms of new teaching that the pilot programme was funding. The questionnaire included a budgetary report as well, with the final page of the questionnaire dedicated to this. The total length of the questionnaire, with all the fields empty, is seven pages, of which supplementary information and budgetary data take up roughly 2½ pages, leaving the rest for the actual open-ended questions regarding the pilot programme (OPH Väliselvitys). This kind of a questionnaire, although short in its length, is still rather demanding and time-consuming due to the nature of the questions, and the data gathered.

The analysis here focuses specifically on the 12th question, that asks regarding differences seen between the genders, and its influence on teaching in the programme. As this is an interim report, as well as since the researchers used open-ended questions, they probably had some uncertainty regarding the type of answers the researchers would receive from the respondents, as open-ended questions are typically used when the researcher needs more information to formulate a proper question in a future questionnaire, or they are uncertain of the types of answers that they will get. This is particularly relevant as the issue of gender in education has seen increased discussion in academic and educational discussion in Finland. The topic remains controversial, and as was discussed in the systematic review of literature's review of Krkovic et al. in "Teacher Evaluation of Student Ability"—which showed that there was no influence on evaluation based on the teachers' gender, with both gendered teachers preferring girls—and Mullola et al.'s "Same Temperament, Different Appreciation"—in which there were significant differences in how teachers evaluated students of the same and different genders—with the contradictory nature of the results from these two studies showing how teachers attitudes towards the topic can vary greatly, and most importantly, how there is a lack of awareness regarding the nature and scope of the issue of gender in educational context with teachers. Despite the high levels of training and education required for teachers in Finland, this lack of awareness will be clearly shown in the answers to the questionnaire that are discussed in the following section.

Further differences between questionnaires are formed based on the types of questions and the kinds of responses allowed for in the questionnaire. The questions themselves range from agree/disagree

semantic differential scales (Cohen et al. 387), a rating scale, to open ended questions that allow the respondent to formulate their own answer (Cohen et al. 382). The choice of question-and-answer format can be determined by the research question, as different question types will produce data that are applicable to specific kinds of research questions. For example, if one wanted to measure the impact of video games on student language skills, a study that asks the respondent to select the appropriate Likert-style grade that reflects their time spent playing, as well as the student's grade in the subject would allow for a relatively easy analysis of possible correlation. The opposite is, naturally, true as well, as open-ended questions enable the researchers to gather more nuanced data, but simultaneously making the analysis of the results more difficult (Cohen et al. 382). The advantage, however, is that this kind of data allows for future analysis and new research to be performed using the resulting data to point a direction.

The scale of data looked for in a questionnaire further determines the types of questions used. This desired scale will then determine the type of question that will be used to gather data of that specific level (Cohen et al. 383). There are advantages and disadvantages to the different types of questions, as Cohen et al. notes in their discussion of dichotomous questions: "[it] is useful, for it compels respondents to 'come off the fence' on an issue" (383). While this kind of a question is very simple and can give great nominal data, it is crucial to consider if the question is appropriate for a simple 'yes/no' answer. The applicability of the type of question must be considered in the context of the study, the participants and the type of data that is being looked for. This limitation and function of the different types of questions is one that follows every type of question used in questionnaires and must always be considered carefully. Choosing between dichotomous, multiple choice, rank ordering, rating scales, constant sum, ratio data and open-ended questions require the specific form of the question to be carefully designed, as they all have their own intricacies and nuances (Cohen et al. 383).

There are further considerations regarding the nature of data a question asks for. Sensitive questions need be formulated with the nature of the question in mind, as there is a difference in respondents answers based on how a question regarding, say, sexual violence is asked (Cohen et al. 395). This issue is further complicated by the difficulty involved with writing a question in general, as language is ambiguous, and understanding is subjective to each individual reading the question (Cohen et al. 396). The number of variables and concerns regarding the specific wording of the questions is vast, and when designing a questionnaire, Cohen et al recommend referring to a source on how to design such a study (395-397). The topic of gender in educational context, while not as emotionally charged as questions of sexual violence, are nonetheless ethically and legally relevant

to teachers. The implication of a difference in treatment or outcome clearly has direct relevance to those who are responsible for the everyday carrying out of education. The answers discussed in the following section show some examples of outright denial of the premise of the question, and some respondents even went as far as to provide counterevidence to show the lack of gendered effect in their schools. These answers and their implications are analysed further in the discussion section.

As human beings are emotional creatures, the order of the questions in a questionnaire can have an impact on the way the respondent answers the questions given (Cohen et al. 397). To reduce the impact of sudden, personal questions, there should be a progress in the types of questions asked, from general, factual questions, towards closed questions, and finishing with open-ended questions (Cohen et al. 398). In the intermediate report survey for the initiative by the Finnish National Agency for Education, there was an open-ended question regarding gendered differences in language learning, and to better analyse the resulting data, the nature of open-ended questionnaires will be discussed next.

The function of open-ended questions in a questionnaire is to allow for a thorough and freestyle answer to a question provided (Cohen et al. 392). This specifically contrasts with closed questions in questionnaires, where even if the respondent were desirous of giving a different answer than the options provided, they are unable to do so. This does not mean that there is no structure or direction given in the question, but rather that the question itself and the way an answer is given are formulated in a manner that encourages expansive, discussive and possibly unexpected or unknown answers. When compared to more closed questions, this possibility of unknown data is especially significant, as one of the main challenges in designing a ranking scale, or dichotomous questions is the exclusionary nature of the question: the possible answers limit the data collected.

This ability to include unexpected data does present some challenges, specifically, in relation to the analysis of the results given (Cohen et al. 393). Word-based data does not translate into statistical analysis easily, as categorizing the answers into discrete groups and then using statistical tools on the results would be questionable as this kind of data is not aggregable (Cohen et al. 393). This issue is compounded by the lack of limitation in the form and length of answers, as comparing two different answers to each other is difficult when the amount of detail, time and effort spent is not necessarily equal, or that the respondents are equally able to "articulate their thoughts and committing them to paper" (Cohen et al. 393). The answers analysed here will be coded, with a grouping based on their semantic content, as it relates to the research question of this section.

Another consideration is the tiring effect of open-ended questions (Cohen et al. 382). Depending on the specific questionnaires format, the questions can be either too vague, causing the respondent to give incorrect information, or to feel that they are unable to answer the question. If vague questions are combined with many open-ended questions, or if the question asks for a large quantity of data, the respondent might simply leave the question unanswered (Cohen et al. 393). This consideration is certainly relevant to the data analysed here, as the questionnaire was demanding, and the respondents' fatigue is a necessary consideration in analysing their answers. On the other hand, this was an interim report sent by the National Agency for Education of Finland, as a part of participating in the pilot programme; the respondents were not "volunteers" in the traditional sense of questionnaire analysis, and as such, the lack of answers in the data should not be assumed to be an indication of fatigue in the manner that it would be in a typical questionnaire study.

3.3 Presenting the data

The questionnaire was administered by the Finnish National Agency for Education, including data collection. During the process of writing this thesis, the initiative manager was contacted for other studies that have been done using data from the interim report, but as of the time of writing, they did not respond to the inquiry. The data used in this analysis, regarding the answers to question 12 of the questionnaire, was sent as an excel file, with no personally identifiable information regarding the respondents. The data was received after reaching out to the manager of the initiative, as well as going through the proper channels at the Agency to approve the distribution of said material.

The total sum of all answers for the questionnaire is 81, and as the question was open-ended, the length of answers varied greatly, with the shortest answer left empty, and the longest answer that was 113 words. The answers varied in language, with six of the answers written in Swedish. These were translated and checked for validity by two others to ensure a proper translation. The answers were further aggregated by coding them into semantic content groups in the following section.

3.4 Coding

The first step in analysing the results of our question is to code the answers based on their results (Cohen et al. 407). A code is "simply a name or label that the researcher gives to a piece of text that contains an idea or a piece of information" (Cohen et al. 559). The initial level of coding is open coding, and is used to categorize the data to groupings, to perform a "data reduction" (Cohen et al. 559, 561). This is especially important in a dataset like the one here, as open-ended questions result in data that is difficult to directly analyse and compare due to the large variance in answers given. The next step in coding can be analytic coding, which aims to interpret the possible reasons of open code categories, i.e., in the case of the questionnaire from the interim report, an analytic code would aim to explore why some answers were left in the "null" open category. This kind of analysis, however, is very challenging and limited in its applicability, as there can only be assumptions as to why a respondent left an answer empty (Cohen et al. 561). After analytic and open coding comes axial coding. This is the grouping of open code categories into related, larger groups of answers based on some shared attribute (Cohen et al. 561-562). The highest level of abstraction for codes is called selective coding, as this category integrates the other types of coding to "form a theory" (Cohen et al. 562). These categories form a hierarchy, with open and descriptive codes being the lowest level, analytical and axial codes in the middle, and selective codes forming the highest order of codes (Cohen et al. 562).

As the data is answering this question: "Have you noticed a difference between the genders in language study and learning? What differences have you noticed?", the answers will be coded into four descriptive codes using the possible answer types: affirmative, negative, contradictory and null. This will be only an initial sorting, as the open-ended nature of the question requires more specific analysis. Dividing the answers in this manner allows to see an overview of the responses for an initial frequency presentation. Typically, an open-ended question will have its answers coded based on a "frequency tally of the range of responses" from a sample of the answers (Cohen et al. 407), but as number of data points is relatively small, the answers are directly categorized rather than sampled.

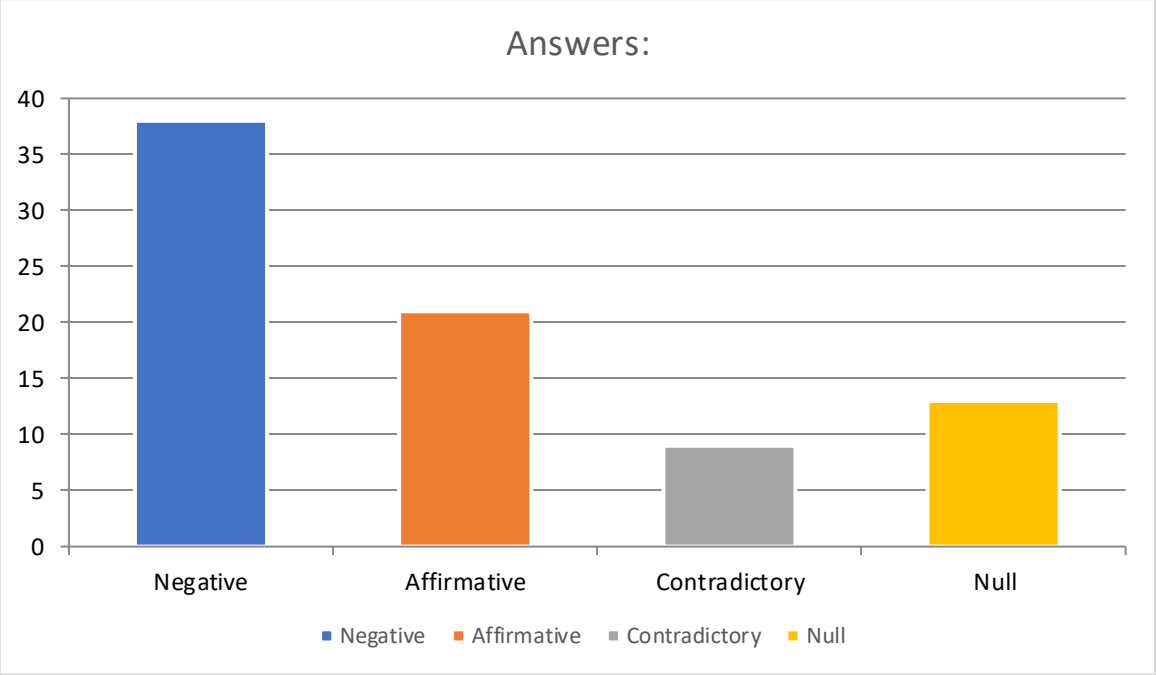


Figure 4

After the initial coding, there were 13 null answers, 38 negative answers, 9 contradictory answers, and 21 affirmative answers. These categories are based on whether the respondents found differences between the genders in language studies or not, with contradictory answers being answers that started by saying that they did not find differences, but then elaborating on differences they found. Null answers were answers that either were completely missing, or did not answer anything relating to the question, for example "question not studied" or simply a dash marker in place of an answer.

3.5 Null answers

These answers are the simplest to describe, as their content allows for no further discussion, while at the same time the most difficult to analyse, as they do not provide an opportunity to understand what the respondents were thinking. The surprising aspect regarding these answers, however, is the fact that 13 out of 81, or 16% of the responses were either left empty, or that the participants in the government programme had not monitored the situation at all.

This lack of participation is even more significant when considering the context of the questionnaire. The report consists of only 15 questions, with a total length of 7 pages regarding the progress of the pilot programme, so the total lack of any answer is rather surprising, as the

respondents could have answered in the negative if they felt that there were no differences between the genders. The possibility that these answers were given with best intentions must also be considered, as the respondents might have felt unable to give an accurate answer to the question. The number of open-ended questions in the questionnaire, as well as the additional data required does increase the likelihood of fatigue causing null answers as well. Unfortunately, this is pure speculation, as all the data available is simply a statement of "not researched" or "not monitored" or simply left empty entirely. The large percentage of answers being in this category will be explored further in the discussion section.

3.6 Negative answers

This was the largest single category of answers with the coding used here, as the differentiation between affirmative and contradictory categories split the possibly affirmative answers into two. There were, however, different codes applied to these answers as well, depending on the answer types:

1. Simple answers without further explanatory or illustrative details
2. Complex answers with additional details

Since the question is open-ended, the length of the answers varied significantly, with the shortest answers being a single word "no", with 5 answers being single words, and with simple negative answers making up most of the negative answers at 27 out of 38. Complex answers were typically greater in length, with the longest answer in this category being 113 words long. The simple answers did have some variance in length as well, but the content of the answer was a negative answer without any additional detail, just with more words, for example:

> (1) "Emme ole havainneet kieltenopiskeluun ja oppimiseen liittyviä eroja sukupuolten välillä," [we have not seen differences between the genders regarding language learning and practice].

Despite the length of the answer, the content meant that these answers were coded into simple negative answers.

The second code of more complex answers allows for a closer study of what the respondents additionally provided in their negative evaluations. These answers made up 11 of the 38 answers in the negative answer category, or 29% of this category of answers. The answers here have a large

variance in the kind of additional information given in the answer, as this was unprompted data that does not directly relate to the question asked in the questionnaire. A common theme amongst all the answers was a focus on personality of the students, rather than gender:

(2) "Sekä tyttöjen että poikien keskuudessa on yhtä lailla kieltenopiskelusta kiinnostuneita ja vähemmän kiinnostuneita oppilaita. Samoin on oppimisen suhteen. Molemmista ryhmistä löytyy kielellisiä yhtäläisyyksiä ja eroja nokkelasti oivaltavia oppijoita. Oppilaan persoona ja kiinnostuksen suuntautumisen kohteet sekä kokemus eri kielistä ovat todennäköisesti kielen omaksumisen kannalta merkittävämpi tekijä kuin sukupuoli." [Among both girls and boys, there are students who are more interested in language learning, and those who are less interested. The same is true for studying in general. Both groups have linguistic similarities and differences learners with clever insights. The personality and interests of the learner, as well as their experience of different languages, are probably more important than gender in determining language acquisition.]

All of these complex negative answers discussed possible reasons for a difference in outcomes that are not based on gender. Of note is one answer that referenced a study from Sweden that looked at how a Steiner Waldorf pre-school could be made into a "genderless" place, as a reason for the lack of a difference between the genders. The presumptive negative answer in this case shows an interesting attitude and approach to the question of gender influence in education from this respondent, as almost a counterargument against the very existence of the question itself in the questionnaire.

3.7 Contradictory answers

This is the smallest category of answers, with only 9 answers making up 11% of the responses. The primary nature of answers here is a direct contradiction in the statements, as in, all the answers start with a statement of either no difference between genders, or that any difference is the result of teaching practices or personality traits. Immediately following this, however, are examples of differences between the genders, for example:

(3) "Ei havaittu suuria eroja sukupuolten välillä. Tytöt ehkä hivenen ahkerampia läksyntekijöitä, pojat rohkeampia puhujia. Pyrimme yleisesti havannoimaan sekä kieltenopiskeluun että muuhun oppimiseen liittyviä eroja henkilöiden kesken, emme

sukupuolten." [We have not noticed large differences between the genders. Girls might be slightly more diligent in doing homework, and boys are braver in spoken language use. We typically try to observe the differences in language learning as well as learning in general between individuals, not genders.]

(4) "Työtapoja tai sisältöjä ei ole ollut tarvetta eritellä sukupuolen mukaan, tämä ei ole ollut ongelma. Pojat ovat uteliaampia ja rohkeampia käyttämään kieltä. Oppiminen on hyvin yksilöllistä, eikä se ole mitenkään sidoksissa sukupuoleen. Oppimiseen vaikuttavat enemmän motivaatio, ympäristö, persoonallisuus ja persoonaan liittyvät tekijät." [Teaching methods or content has not been differentiated between genders, this has not been a problem. Boys are more curious and braver in their use of language. Learning is a very individual process, and is in no way tied to gender. Motivation, environment, personality and individual characteristics affect learning.]

These answers were always with an initial "no" answer, as any answers that started with an agreeing statement are included in the affirmative answers. The lack of any answers that would claim a difference, but then showing similarities is striking, as there is no inherent reason for the lack of these kinds of answers. In another example, a participating school had four students drop out of the pilot initiative's early language courses, three of whom were boys and one of whom was a girl. The girl dropped the programme due to a parent's choice, because they wanted the student to go and learn another language instead, and the pilot programme was seen as a complicating factor in this endeavour. The other three, all of whom were boys, were categorized as having trouble focusing or lacking motivation. Oddly, the respondent did not see any gender effect in this case, even though three out of four students who stopped were boys, and the reasons given for them stopping participation in the pilot program were negative traits rather than conflicting schedules. The fact that the respondent did not see a conflict based on gender here is significant, as studies presented in the systematic review section of this thesis show differences between the genders in language learning, teacher evaluations and perceived student behaviour patterns (See Krkovic et al.; Lepola; Mullola et al. for examples).

Like the negative complex answers, all the answers in this category share the theme of personality being the major factor in the differences between students. Five of the answers noted that boys prefer more active learning practice, and two mentioned video games as an influence on vocabulary of boys. Girls were seen as being more conscientious with homework, and more able to do longer, quieter practice forms, as well as being more likely to select a foreign language as an elective.

The apparent contradictions found within this short, open-ended question's answers require careful consideration. All the respondents are at minimum master's degree holders, and professional teachers in Finland. The fact that the responses are so clearly contradictory in their content seems to indicate a significant framing issue around the respondents view of gender, or an inability to see how it influences educational outcomes. The analysis section will consider this question in more detail.

3.8 Affirmative answers

The answers in this category are the second largest group, at 21 out of 81, or 26% of the answers. Certain common themes appear frequently here as well, in the same way that they did in the other categories. Of the 21 answers, eight specifically mention that boys have better vocabulary and language skills due to video games, for example:

(5) "Pelaaminen näkyy pojilla jo omaksutun sanavaraston laajuudessa ja valmiudessa ymmärtää puhuttua englantia." [Gaming already shows in the boys' vocabulary size and their preparedness to understand spoken English]

The attribution of the differences between the genders followed a relatively repetitive format, with boys' skills typically attributed to playing video games, and girls' skills seen as the results of effort. Additionally, boys were seen as "braver" users of language, and girls as more perfectionists, with boys preferring physically active learning, and girls performing better in traditional teaching. There were a few outliers, one answer specifically wrote that girls were more likely to participate in class activities such as singing, and boys being quieter, with another respondent saying that girls were better able to show their skills in class. These cases stand out as the exception from the other responses, as they were the only ones in which girls' behaviour was characterized as more active, and boys' as quieter. Finally, one respondent simply stated that of those participating in the pilot programme, two thirds were girls, and one third were boys. Some examples of the responses are:

(6) "Tytöt ovat rohkeampia tekemään/osallistumaan/laulamaan erilaisissa tilanteissa ja pojat enemmän tarkkailevat, mutta ovat kuitenkin aktiivisesti läsnä. Molemmat kuitenkin tuntuvat oppivan." [Girls are braver in doing/participating/singing in different types of situations, and boys are more likely to observe, but are still actively present. Both do appear to be learning nonetheless.]

(7) “Asenteissa, opiskelutyyleissä ja pohjatiedoissa on jonkin verran eroja tyttöjen ja poikien välillä, joskin yksilölliset erot ovat sukupuolieroja suurempia. Erityisesti poikien opiskelussa näkyy usein kokemus suullisen kielen käytöstä esimerkiksi pelimaailmassa ja näin ollen monilla on hyvä pohja kielten opiskelulle. Tämä lisää myös motivaatiota kielten, erityisesti englannin, opiskeluun ja tekee siitä helpompaa.” [There are some differences between the genders in their attitude, learning styles, and foundational knowledge, although individual differences are larger than gendered ones. Experience with spoken language use is especially noticeable in boys, for example from gaming, and many have a good foundation for learning from this. This also increases motivation in language studies, especially with English, and makes it easier.]

The answers in this category reflect common themes found in both the literature review, and in the other categories of answers: boys are seen as active, less able to focus and be quiet, while girls are seen as perfectionists and are better able to focus on quiet work. While the data presented seems to give a strong case for how things are, the answers are what the respondents perceived the differences to be, they do not give any empirical data about the actual situation in class. The questionnaire can only give us data as to what the respondents are willing to answer in each situation, not as to what is the “actual” situation in the world.

3.9 Discussion

The data provided by the National Agency for Education allows for a glimpse into the perceptions of experienced teachers in Finland, specifically as to how they view and perceive gender differences in L2 learning between boys and girls at school. The pilot initiative specifically focused on lowering the starting age of early L2 learning, and as such, any possible differences found between the genders is of even higher importance, as most of the students taught in the initiative would in elementary school education. The fact that these differences were found in pupils this young is a sign of the extent of the difference found in the Finnish educational system.

That aspect, however, is better covered in the studies reviewed in the systematic review of literature section of this thesis. The larger implication regarding the data analysed in this section is the apparent disconnect between teachers’ perceptions around and the reality of the effect of gender on

education. Most answers are either null (13 out of 81) or negative (38 out of 81), or put in another way, 63% of the respondents either did not think the question relevant, did not have a chance to study it, did not fill in the question, or found no differences between the genders. This is especially concerning as shown in the introduction section of this thesis, the gender gap in educational outcomes in Finland has been well past the equivalence point for about 20 years now, and the gap is increasing in size each year. The most charitable explanation of this disconnect would be to say that there is no difference at all at the specific schools that participated in the pilot programme, and that the large statistical differences found in educational outcomes are due to structural issues in how high school and university programmes are selecting their students. The selection process of secondary and tertiary education programmes is certainly a factor in the overall effect that is causing the large gap in educational outcomes. It seems unlikely that in these 51 schools who participated in the pilot programme there is no gendered effect, when it is seen in almost all OECD countries.

The more likely explanation is that the effect of gender is relatively small at the individual level, overshadowed by personality traits and environmental factors. This was also seen in practically all the answers analysed here. The most common reason suggested to explain boys' better performance in English was given as their proclivity towards video games, or how girls' ability to follow a traditional teaching session was seen as better. In the contradictory answer category, the regularity with which these explanatory statements were used was very high, with four of the nine contradictory answers specifically mentioning personality as the reason for the differences found. Affirmative answers had no such explanatory statements, as the findings simply expand upon what differences between the genders the respondents had seen. The implication of these figures is, however, rather stark, as when we add up the number of null, negative, and contradictory answers that specifically exclude gender as a possible reason, the total number is 55 respondents out of 81. Put in another way, 68% of the respondents did not consider gender to be a valid reason for a difference in educational outcome.

While characterizing the responses in this way might be somewhat inflammatory, the numbers do not exaggerate. There is a possibility that there truly is no difference between the genders in these 55 schools' programmes that excluded gender as an influencing factor, but this does seem rather unlikely. The relevance of this specific lack of teacher awareness on gender is significant with regards to the research question of this thesis, namely, trying to find explanations for the large and growing gender gap in educational outcomes in Finland. While most of the research articles included in the systematic review section of this thesis focused on students' performances, the two

studies mentioned repeatedly are crucial in understanding the context of the results from this questionnaire. Krkovic et al. showed that even when controlling for objective performance, the teachers in that study overestimated girls' performance in school. The answers discussed above regarding students who dropped out of the initiative lessons, who happened to be three boys and one girl, seems to fit into the study's context perfectly, as the girl stopped attending due to schedule conflicts, but the boys were attributed to inability to focus and do the work. The respondent in that case specifically said that there was no difference between the genders, and that both boys and girls have highly motivated and bored students equally, seemingly unaware of the perfect case of gendered difference in their own answer. The topic of teacher perceptions and temperament was researched in Mullola et al., and that paper found that there were great differences in how teachers evaluated students of the opposite gender, and how they perceived their students' teachability. This was closely tied to their temperament and behaviour in class. The fact that so many respondents in this questionnaire felt justified in their exclusion of gender as an affecting factor in educational outcomes should be of grave concern, as based on the evidence shown in the first section this thesis the effect is very real.

4 Conclusion

The ever-increasing scale of the gap between men's and women's attainment in education is clearly seen from the statistics collected by Statistics Finland, with 30 percent more female than male masters' degree graduates in 2019. The difference is seen at other levels of education as well, and men are more likely to drop out of education than women. However, as discussed in this thesis, only looking at statistics and graduation rates does not allow for causal analysis, nor does it elucidate possible confounding factors in the realities that result in the statistics. While there is an apparent advantage to girls' educational performance, the pay gap still exists, there are many fields in academia and sciences that are male dominated despite the higher number of female PhD graduates, and the systemic nature of sexism is still very much in place even in as "advanced" of a society as Finland.

There is a lack of a general discussion in academia, education, and society at large around the changing role of men in this context. The role of women has been under concerted, deliberate, and extensive discussion for the last hundred years at least, with a focus on emancipating women. The

results are evident: achievements include voting rights, divorce rights, participation in the work force and targeted programmes to enable women in male dominated. Such a discussion is, however, missing for the role of men. The gender roles are changing rapidly in society at large, with more and more men staying at home, being the caregivers of a family's children, and not participating in the workforce. Even when men do participate as workers, the concerted drive to have more inclusive workplaces means that traditional gender norms are being challenged every day. Specific to Finland, the existence of mandatory military service applies only to men, but there seemingly are no studies that research the effect that taking a year's break from school has on educational attainment, nor is there any discussion about how military service affects the gender norms of those who must participate in it by law, or generally their behaviour. Simultaneously, there is no great academic discussion in the vein of Simone de Beauvoir, Betty Friedan or Naomi Wolfe for men, the so-called men's rights movement is mostly co-opted by misogynists and those who feel threatened by the equality strived for by feminism. There are many legitimate, important issues around men that are effectively being ignored, and the need for a larger discussion about men and their changing roles is desperate.

As the topic of the gender gap in educational attainment topic is of incredible complexity and has many variables that affect it, the scope of this thesis is not to try and find a cumulative, total summary of the effect of gender in education. Instead, the first section of this thesis showed the trends in education that have been continuing in the Finnish education system for the past 45 years. In the systematic review of literature, the current best research on the topic was reviewed and discussed. The near universality of findings in contemporary research literature is a clear indication that this effect is not a matter of opinion or perception, studies ranging from pre-natal babies (Korpilahti, Kaljonen and Jansson-Verkasalo) to end of middle school students (Pesu et al., Torppa et al., Manu et al.) all showed either equivalency between the genders, or an advantage to girls. Moreover, Orellana et al., along with the PISA scores from 2018 show that this problem is not specific to Finland, it exists in all developed countries to varying degrees.

The specific differences found in the systematic review between genders were quite varied, with a general focus on language skills (Mäki et al., Korpilahti et al., Orellana et al., Peura et al., Pesu et al., Torppa et al.) being very common. These studies included examples of how boys and girls engaged in different kinds of reading, with girls reading more and boys reading more digital texts than girls, different levels of motivation towards reading, as well as different leisure time reading habits. The surprising result of the systematic review, however, was the lack of extensive studies of this phenomenon. The Boolean search resulted in 124 possible studies, of which 29 were selected

for close reading. From these, only 15 were applicable to the study at hand, and some of the resulting ones could be argued to only be tangentially relevant, such as the study on eating habits by de Oliveira Figueiredo, et al., where the study may allow for an inference to be made between school performance and eating habits of students, but the study itself did not make this link. The lack of targeted, large-scale studies regarding the educational gap is surprising, especially when considered in the context of wider societal trends that were examined in the first section of this thesis. There is a great need for research on this topic, specifically aimed at determining the causes of the expanding gender gap over the whole population, rather than looking at specific, smaller influencing factors.

The differences between the genders continued in temperament, behaviour, and task-orientation, with Lepola showing pre-schooler behavioural differences, and Mullola et al. discussing ideal student temperament and behavioural modes. Pöysä et al. studied the effects of engagement on GPA and found that boys showed worse engagement in four categories, and girls had no worse engagement in any category. While there are many different factors that influence educational outcome, the role that teachers play is clearly central. Although the two studies of Krkovic et al. and Mullola et al. clearly show that teachers are influenced by the gender of their students in how they interact with and evaluate them, the aim for educators in an ideal situation would be that there is no such difference, or at the very least, that such differences are minimized. There is no fundamental reason for a gendered effect on teachers work, but the difficulty seems to come from lack of awareness of such bias in working teachers. If there is no awareness of a problem, there can be no way to resolve it. Determining the extent and nature of an issue is the first step in working towards a solution. It was for this reason that the analysis of the interim report answer is so crucial, to show an example of current teacher perceptions around this issue.

This was done in the final part of the thesis, where an analysis of a recent interim report of a government pilot programme to bring L2 education into earlier grades was made, with a question in this interim report specifically asking about what, if any, differences had the respondents found between the genders. Questionnaires are a commonly used tool in researching attitudes and perceptions. In the case of this specific question and questionnaire, there was a chance to measure the perceptions and attitudes of Finnish educators as they relate to a very consequential and timely topic, that of gender in educational outcomes. The results are troubling, as almost 70% of the respondents either did not see any differences, attributed the differences found to personality rather than gender, did not monitor their students in this way, or simply did not answer the question at all.

As the data, statistics, and research in this thesis has shown, there is a large, persistent, and growing gap between boys and girls' educational performance in Finland. Studies discussed in this thesis have shown that teachers view boys as more active, more rebellious and associate them with fewer attributes of an "ideal student". Boys are more likely to fail a class, to be held back a year, to drop out and to not enter higher education. The number of factors that influence these statistics are clearly very high, and to attempt to point to a simple causal factor is impossible. This topic urgently needs targeted, large-scale study if we are to reverse the current trends, as the current Finnish education system has a clear and systemic bias that enables girl students to perform better.

Works Cited

"Best Countries for Women" *US News and World Report,* https://www.usnews.com/news/best-countries/best-women. Accessed 8 April 2021.

Cohen, Louis, et al. *Research Methods in Education*. Vol. 7th ed, Routledge, 2011. *EBSCOhost*, search-ebsco-host-com.pc124152.oulu.fi:9443/login.aspx?direct=true&db=nlebk&AN=548475&site=ehost-live&scope=site.

de Oliveira Figueiredo, Rejane Augusta, et al. "Identifying Eating Habits in Finnish Children: A Cross-Sectional Study." *BMC public health* 19.1 (2019): 312. Web.

Fröjd, S. A., et al. "Depression and School Performance in Middle Adolescent Boys and Girls." *Journal of adolescence* 31.4 (2008): 485-98. *SCOPUS.* Web. 12 March 2021.

OECD (2019), *Finland – Country Note – PISA 2018 Results*, PISA, OECD Publishing, Paris, https://www.oecd.org/pisa/publications/PISA2018_CN_FIN.pdf.

Accessed 8 April 2021.

OECD, *PISA 2018 Results – Combined Executive Summaries Volume I, II & III*, PISA, OECD Publishing, Paris, https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf. Accessed 8 April 2021.

OECD, *PISA FAQ,* OECD.org, https://www.oecd.org/pisa/pisafaq/. Accessed 8 April 2021.

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

OECD (2020), "Girls' and boys' performance in PISA", in PISA 2018 Results (Volume II): Where All Students Can Succeed, OECD Publishing, Paris. DOI: https://doi.org/10.1787/f56f8c26-en

Official Statistics of Finland (OSF): Discontinuation of education [e-publication].
ISSN=1798-9302. 2017, Appendix table 3. Comprehensive school drop-outs in academic years 1999/2000 to 2017/2018. Helsinki: Statistics Finland [referred: 8.4.2021].
Access method: http://www.stat.fi/til/kkesk/2017/kkesk_2017_2019-03-14_tau_003_en.html

Opetushallitus (OPH), *Opettajat ja rehtorit Suomessa 2019: Esi- ja perusopetuksen opetusryhmät*, web, accessed 05.06.2021, https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/opettajat-ja-rehtorit-suomessa-2019-esi-ja-perusopetuksen

Papadopoulos, Anna "RANKED: World's Best Countries for Women, 2020" *CEOWorld Magazine*, 21 June 2020, https://ceoworld.biz/2020/06/21/ranked-worlds-best-countries-for-women-2020/

Korpilahti, P., A. Kaljonen, and E. Jansson-Verkasalo. "Identification of Biological and Environmental Risk Factors for Language Delay: The Let's Talk STEPS Study." *Infant Behavior and Development* 42 (2016): 27-35. *SCOPUS.* Web. 12 March 2021.

Krkovic, K., et al. "Teacher Evaluation of Student Ability: What Roles do Teacher Gender, Student Gender, and their Interaction Play?" *Educational Research* 56.2 (2014): 244-57. *SCOPUS.* Web. 12 March 2021.

Lepola, J. "The Role of Gender and Reading Competence in the Development of Motivational Orientations from Kindergarten to Grade 1." *Early Education and Development* 15.2 (2004): 215-40. *SCOPUS.* Web. 12 March 2021.

Lisbeth M. Brevik, Glenn Ole Hellekjær, Outliers: Upper secondary school students who read better in the L2 than in L1, International Journal of Educational Research, Volume 89, 2018, Pages 80-91, ISSN 0883-0355, https://doi.org/10.1016/j.ijer.2017.10.001. (https://www.sciencedirect.com/science/article/pii/S0883035517310480)

Manu, M., et al. "Kindergarten Pre-Reading Skills Predict Grade 9 Reading Comprehension (PISA Reading) but Fail to Explain Gender Difference." *Reading and Writing* 34.3 (2021): 753-71. *SCOPUS.* Web. 12 March 2021.

Mullola, S., et al. "Same Temperament, Different Appreciation: Temperament's Influence on School Achievement Depends on Student's and Teacher's Gender and Teacher's Age." *Psychology of Gender Differences.*, 2012. 77-105. *SCOPUS.* Web. 12 March 2021.

Mäki, H. S., et al. "Predicting Writing Skill Development with Word Recognition and Preschool Readiness Skills." *Reading and Writing* 14.7-8 (2001): 643-72. *SCOPUS.* Web. 12 March 2021.

Statistics Finland. "Educational Structure of Population 1970-2019." Web. 8 April 2021

Opetushallitus, *Hakutiedote 2018 – Kärkihanke 1: Uusi peruskoulu – ohjelma, kieltenopetusta koskeva koulu*. OPH-228-2018.

Opetushallitus, *Väliselvitys 2018 - Valtionavustukset kielten opiskelun varhentamiseen, kehittämiseen ja lisäämiseen - Hakuryhmä A*, Retrieved 02.12.2021, Web, https://valtionavustukset.oph.fi/avustushaku/140/valiselvitys

Oppermann, E., et al. "Elementary School Students' Motivational Profiles Across Finnish Language, Mathematics and Science: Longitudinal Trajectories, Gender Differences and STEM Aspirations." *Contemporary educational psychology* 64 (2021) *SCOPUS.* Web. 12 March 2021.

Orellana, P., et al. "Motivation to Read in Grades K–2: A Cross-Cultural Perspective." *International Journal of Early Years Education* 27.4 (2019): 423-40. *SCOPUS.* Web. 12 March 2021.

Pesu, Laura, et al. "The Development of Adolescents' Self-Concept of Ability through Grades 7-9 and the Role of Parental Beliefs." *Frontline Learning Research* 4.3 (2016): 92-109. Web.

Peura, P. I., et al. "Specificity of Reading Self-Efficacy among Primary School Children." *Journal of Experimental Education* 87.3 (2019): 496-516. *SCOPUS.* Web. 12 March 2021.

Pöysä, S., et al. "Adolescents' Engagement Profiles and their Association with Academic Performance and Situational Engagement." *Learning and Individual Differences* 82 (2020) *SCOPUS.* Web. 12 March 2021.

Silinskas, G., et al. "Maternal Teaching of Reading and Children's Reading Skills in Grade 1: Patterns and Predictors of Positive and Negative Associations." *Learning and Individual Differences* 27 (2013): 54-66. *SCOPUS.* Web. 12 March 2021.

Sol Lim, Cheol E. Han, Peter J. Uhlhaas, Marcus Kaiser, Preferential Detachment During Human Brain Development: Age- and Sex-Specific Structural Connectivity in Diffusion Tensor Imaging (DTI) Data, *Cerebral Cortex*, Volume 25, Issue 6, June 2015, Pages 1477–1489, https://doi.org/10.1093/cercor/bht333

Torppa, Minna, et al. "Why do Boys and Girls Perform Differently on PISA Reading in Finland? the Effects of Reading Fluency, Achievement Behaviour, Leisure Reading and Homework Activity." *Journal of Research in Reading* 41.1 (2018): 122-39. Web.

World Economic Forum, *Global Gender Gap Report 2020*, 2019, http://www3.weforum.org/docs/WEF_GGGR_2020.pdf. Accessed 8 April 2021.