



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

THE GENDER EQUALITY PARADOX IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS EDUCATION: A FOCUS ON INTRAINDIVIDUAL ACADEMIC STRENGTHS

Marco Balducci



TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

THE GENDER EQUALITY PARADOX IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS EDUCATION: A FOCUS ON INTRAINDIVIDUAL ACADEMIC STRENGTHS

Marco Balducci

University of Turku

Faculty of Social Sciences
Department of Psychology and Speech-Language Pathology
Subject: Psychology
Doctoral Programme in Inequalities, Interventions, and New Welfare State

Supervised by

Professor, Elina Kilpi-Jakonen
University of Turku

Associate Professor, Karin Halldén
Stockholm University

Senior Researcher, Marie-Pier Larose
University of Turku

Reviewed by

Associate Professor, David Giofrè
University of Genova

Associate Professor, Jonathan Wai
University of Arkansas

Opponent

Associate Professor, David Giofrè
University of Genova

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-952-02-0215-6 (PRINT)
ISBN 978-952-02-0216-3 (PDF)
ISSN 0082-6987 (Print)
ISSN 2343-3191 (Online)
Painosalama, Turku, Finland 2025

*To Dylan, Noemi, and Aurora—
my lights in the darkness and the reasons I never gave up.*

UNIVERSITY OF TURKU

Faculty of Social Sciences

Department of Psychology and Speech-Language Pathology

Subject: Psychology

MARCO BALDUCCI: The Gender Equality Paradox in Science,

Technology, Engineering, and Mathematics Education: A Focus on

Intraindividual Academic Strengths

Doctoral Dissertation, 147 pp.

Doctoral Programme in Inequalities, Interventions and New Welfare State

June 2025

ABSTRACT

Whether sex differences increase or decrease as more gender equality is achieved remains debatable. Some scholars argue that these differences result from socialization processes and, thus, should diminish when countries progress toward gender equality. Conversely, others predict the opposite pattern, which they refer to as the gender equality paradox. The debate has primarily focused on sex disparities in inorganic science, technology, engineering, and mathematics (STEM) educational fields due to their association with various long-term socioeconomic outcomes, such as the gender pay gap. In fact, despite numerous efforts, women remain underrepresented in these fields—a trend that seems particularly pronounced in more gender-equal countries.

Among the numerous STEM-related factors (e.g., academic skills and personality) proposed to influence this paradoxical trend, intraindividual strengths have recently emerged as a potential critical component. Intraindividual strengths refer to a student's relative advantage in mathematics, reading, or science compared with their overall scores and independent of peers' performance. Girls generally perform better in reading than in mathematics or science, while boys generally perform better in mathematics or science than in reading. Sex differences in intraindividual strengths have been linked to career choices in STEM and also appear to follow a pattern consistent with the gender equality paradox.

However, there has not yet been a comprehensive assessment of the gender equality paradox in STEM-related factors, and it remains unclear whether this phenomenon is widespread or specific to a few domains. Moreover, despite their potential relevance, sex differences in intraindividual academic strengths remain largely understudied.

This study aimed to synthesize the literature on the gender equality paradox in STEM-related factors, such as academic skills and personality, to develop a theoretical framework for future research addressing this phenomenon. Additionally, it sought to examine sex differences in intraindividual academic strengths to shed light on the underrepresentation of women in STEM fields.

The research comprised three studies. Following the PRISMA guidelines, Study I involved a systematic review of published scientific literature on STEM-related sex differences in academic skills (in mathematics, reading, and science), as well as

personality models (Big Five, HEXACO, basic human values, and vocational interests), and their associations with gender equality indicators. Subsequently, employing data from five recent waves (2006–2018) of the Programme for International Student Assessment (PISA), along with Spearman correlation and ordinary least squares models, Studies II and III investigated how sex differences in intraindividual academic strengths vary with increases in the Global Gender Gap Index (GGGI) from the World Economic Forum—both overall (Study II) and at different levels of academic achievement (bottom 5%, top 5%, and average) (Study III).

The literature review indicates that the gender equality paradox has been extensively studied across many domains and is consistently observed, particularly in personality traits and reading skills. In contrast, sex differences in mathematics appear to emerge regardless of country-level gender equality. Additionally, the review highlighted the potential key role of intraindividual academic strengths in explaining the underrepresentation of women in STEM, although few studies have addressed them. Building on these findings, the results from Studies II and III align with previous research on sex differences in intraindividual strengths. A relative advantage for girls in reading was observed widely across countries and achievement levels, just as a relative advantage for boys was found in mathematics and science. Additionally, the gender equality paradox consistently emerged in sex differences in reading and science as intraindividual strengths, both across PISA waves and achievement levels. In contrast, sex differences in mathematics as intraindividual strengths do not appear to be related to the GGGI.

In the last part of this dissertation, the results are discussed in the context of expectancy-value theory, which posits that intraindividual strengths should influence future career choices. On this premise, limitations are addressed, and recommendations are proposed for future research as well as new policies aimed at reducing the underrepresentation of women in STEM.

KEYWORDS: Science, technology, engineering, and mathematics; STEM; gender equality; intraindividual strengths; personality; sex differences; Programme for International Student Assessment

TURUN YLIOPISTO

Yhteiskuntatieteellinen tiedekunta

Psykologian ja logopedian laitos

Oppiaine: Psykologia

MARCO BALDUCCI: Sukupuolten tasa-arvon paradoksi luonnontieteiden, teknologian, insinööritieteiden ja matematiikan koulutuksessa:

Yksilönsisäisiin akateemisiin vahvuuksiin keskittyvä tarkastelu

Väitöskirja, 147 s.

Eriarvoisuuden, interventioiden ja hyvinvointivaltion tutkimuksen tohtoriohjelma

Kesäkuu 2025

TIIVISTELMÄ

On edelleen kiistanalaista, lisääntyvätkö vai vähenevätkö sukupuolierot sukupuolten tasa-arvon edistyessä. Osa tutkijoista väittää, että nämä erot johtuvat sosiaalisuomisprosesseista ja siten niiden tulisi vähentyä maiden tasa-arvoistuesssa. Toiset taas ennustavat päinvastaista ilmiötä, jota kutsutaan sukupuolten tasa-arvon paradoksiksi. Keskustelu on keskittynyt erityisesti sukupuolieroihin epäorgaanisilla STEM-aloilla (luonnontieteet, teknologia, insinööritieteet ja matematiikka), koska nämä alat ovat yhteydessä lukuisiin pitkäaikaisiin sosioekonomisiin lopputulemiin, kuten sukupuolten väliseen palkkaeroon. Lukuisista ponnisteluista huolimatta naiset ovat edelleen aliedustettuina näillä aloilla, ja tämä suuntaus näyttää olevan erityisen voimakas tasa-arvoisissa maissa

Monien STEM-alaan liittyvien tekijöiden (esim. akateemiset taidot ja persoonallisuuspiirteet) joukossa on viime aikoina noussut esiin uusi mahdollinen keskeinen selittävä tekijä: yksilönsisäiset vahvuudet. Näillä tarkoitetaan oppilaan suhteellista vahvuutta matematiikassa, lukemisessa tai luonnontieteissä hänen omiin kokonaispisteisiinsä verrattuna – siis riippumatta muiden suoriutumisesta. Tytöt suoriutuvat yleisesti paremmin lukemisessa kuin matematiikassa tai luonnontieteissä, kun taas pojat suoriutuvat keskimäärin paremmin matematiikassa ja luonnontieteissä kuin lukemisessa. Sukupuolierot näissä yksilönsisäisissä vahvuuksissa ovat yhteydessä uravalintoihin STEM-aloilla ja näyttävät seuraavan mallia, joka on yhdenmukainen sukupuolten tasa-arvon paradoksin kanssa.

Laaja-alaista arviota sukupuolten tasa-arvon paradoksista STEM-alaan liittyvissä tekijöissä ei kuitenkaan ole vielä tehty, eikä ole selvää, onko ilmiö yleinen vai rajoittuuko se vain muutamaisiin osa-alueisiin. Lisäksi, vaikka näillä yksilönsisäisillä akateemisilla vahvuuksilla voi olla merkittävää selitysvoimaa, niitä ei ole juurikaan tutkittu sukupuolierojen näkökulmasta.

Tämän tutkimuksen tavoitteena oli koota yhteen aiempi kirjallisuus sukupuolten tasa-arvon paradoksista STEM-alaan liittyvissä tekijöissä, kuten akateemisissa taidoissa ja persoonallisuuspiirteissä, ja kehittää teoreettinen viitekehys ilmiön tulevaa tutkimusta varten. Lisäksi tavoitteena oli tarkastella sukupuolieroja yksilönsisäisissä akateemisissa vahvuuksissa ja tuottaa uutta ymmärrystä siitä, miksi naiset ovat edelleen aliedustettuina STEM-aloilla.

Tutkimus koostui kolmesta osatutkimuksesta. Osatutkimus I oli PRISMA-ohjeistuksen mukaisesti tehty systemaattinen kirjallisuuskatsaus julkaistuihin tieteellisiin artikkeleihin, joissa käsitellään sukupuolieroja STEM-alaan liittyvissä akateemisissa taidoissa (matematiikassa, lukemisessa ja luonnontieteissä) sekä persoonallisuummalleissa (Big Five, HEXACO, perustavanlaatuiset arvot ja ammatilliset kiinnostuksen kohteet) ja niiden yhteyttä sukupuolten tasa-arvoa mittaaviin indikaattoreihin. Tämän jälkeen osatutkimuksissa II ja III käytettiin viittä PISA-aineistonkeruuta (2006–2018) sekä Spearmanin korrelaatio- ja pienimmän neliösumman regressiomalleja sukupuolierojen tarkastelemiseksi yksilönsisäisissä akateemisissa vahvuuksissa suhteessa Maailman talousfoorumin globaaliin sukupuolten tasa-arvoindeksiin (Global Gender Gap Index, GGGI) – sekä yleisesti (osatutkimus II) että eri suoriutumistasoilla (alimmat 5 %, ylimmät 5 % ja keskiarvo) (osatutkimus III).

Kirjallisuuskatsaus osoitti, että sukupuolten tasa-arvon paradoksia on tutkittu laajasti eri osa-alueilla ja se toistuu johdonmukaisesti, erityisesti persoonallisuuspiirteissä ja lukutaidoissa. Sen sijaan erot matematiikan taidoissa näyttävät ilmenevän sukupuolten tasa-arvosta riippumatta. Katsauksessa korostui myös yksilönsisäisten akateemisten vahvuuksien mahdollisesti keskeinen rooli naisten aliedustuksen selittämisessä STEM-aloilla, vaikka niitä on tutkittu toistaiseksi vähän.

Näiden löydösten pohjalta osatutkimusten II ja III tulokset tukevat aiempaa tutkimusta sukupuolieroista yksilönsisäisissä vahvuuksissa. Tyttöillä havaittiin olevan suhteellinen etu lukemisessa, ja pojilla vastaavasti matematiikassa ja luonnontieteissä – kaikissa maissa ja eri suoriutumistasoilla. Lisäksi sukupuolten tasa-arvon paradoksi ilmeni johdonmukaisesti sukupuolieroissa lukemisen ja luonnontieteiden yksilönsisäisissä vahvuuksissa sekä eri vuosina kerätyissä PISA-aineistoissa että eri suoriutumistasoilla. Sen sijaan sukupuolierot matematiikan yksilönsisäisissä vahvuuksissa eivät näytä olevan yhteydessä GGGI-indeksiin.

Väitöskirjan yhteenveto-osassa tuloksia tarkastellaan odotusarveteorian (expectancy-value theory) näkökulmasta, jonka mukaan yksilön vahvuudet vaikuttavat uravalintoihin. Tämän pohjalta käsitellään tutkimuksen rajoitteita ja esitetään suosituksia jatkotutkimukselle sekä uusia politiikkatoimia, joiden tavoitteena on vähentää naisten aliedustusta STEM-aloilla.

ASIASANAT: Luonnontieteet, teknologia, insinööritieteet ja matematiikka; STEM; sukupuolten tasa-arvo; yksilönsisäiset vahvuudet; persoonallisuus; sukupuolierot; kansainvälinen oppimistulosten arviointiohjelma (PISA)

Acknowledgments

This thesis represents one of the most meaningful chapters of my life. As with any story, while the protagonist may be at the center, it is the people around him who make the story worth reading. Without them, this would just be another doctoral dissertation.

My supervisor, Professor Elina Kilpi-Jakonen, is the first person without whom this research would not exist. I am deeply grateful for your guidance, your support, and most of all, your patience. Thank you for believing in me, even when it might have been easier not to. You are one of the researchers I admire most, and I hope to one day reach even a fraction of your skills.

This work would not have been possible without INVEST. I would like to thank Professor Jani Erola for creating an environment rich in opportunities for meaningful research. Leadership is often associated just with “power”, but real leadership means creating space for others to grow. In this sense, Jani represents to me what it means to lead with vision and generosity.

Thank you also to my co-supervisor, Associate Professor Karin Halldén, for your feedback and guidance. Your comments helped me stay focused and balanced in my views, and this thesis is stronger because of your contribution.

Professor Jukka Leppänen played an important role in aiding me navigate the wild but fascinating world of academia during the early stages of my Ph.D. I still remember our conversations about the best steps to take, and this achievement is, in part, the result of those early, well-guided decisions.

I am grateful to the Finnish Cultural Foundation, especially the “Huhtamäen rahasto”, for financially supporting this research. In an academic environment where political ideologies can sometimes cloud judgment, I appreciated your support for a project that somehow challenges mainstream views in a scientifically grounded way.

A heartfelt thank you to Distinguished Professor David C. Geary, whom I consider my academic mentor. I began this journey reading your work and never imagined I would one day co-author a paper with you. I have learned so much from our collaboration and from our conversations during my visit to the University of Missouri. I look forward to continuing this work together.

The same goes for Professor Gijsbert Stoet. Though we have not yet met in person, I am truly grateful for your generosity in sharing your knowledge and for the

opportunity to work with you on what has become one of my most important publications so far. I cannot wait for the hopefully many more collaborations to come and to keep learning from you.

I would also like to thank someone who, sadly, is no longer with us—Dr. Lucia Ruggera. You guided me in the early stages of this project and played a key role in shaping its foundation. I will always carry with me what you taught me.

Judith, do you remember everything we went through in Stockholm? That was the hardest period of my Ph.D., and I would not have made it through without your support. Our friendship means a great deal to me, and I hope it will last for many, many years to come. I also hope that you find an institution that values your skills and gives you the opportunities you deserve.

Thank you to my friends and colleagues—Charles, Ryan, Hector, and Lydia. Your friendship kept me grounded and motivated through the many ups and downs of this crazy Ph.D. journey. I am sure there are still many laughs ahead, wherever academia takes us.

I also greatly appreciated Laura Heiskala's help, both in navigating the practicalities of my defense and in proofreading the Finnish abstract of this dissertation. Receiving guidance from someone who had recently completed their own defense made the entire process much smoother.

A special thank you to Waseem Haider. We started this experience together during COVID, often alone with just our research. We have grown a lot, professionally and personally, since then. I know sometimes you had to suffer the consequences of being my friend, and I appreciate your patience and integrity through it all.

Another special thank you goes to my friend, office mate, and unofficial complaint manager, Francesco. You have been a constant source of support and perspective. I cannot count the times you helped me refocus and get back on track when I was too frustrated to see clearly. I still hear your voice in my head whenever I write, pushing me to do better. It is annoying, but in the best, most constructive way.

Lastly, the person most responsible for helping me reach this goal: the one and only Dr. Marie-Pier Larose. I do not know where to begin. Just know that I am endlessly grateful. This is not just my achievement, it is also yours. Without you, I would still be lost somewhere in the midst of statistics. If I can dream of reaching something truly meaningful one day, it is because of you. Your open-mindedness, depth of knowledge, and intellectual integrity are rare in academia. I will never forget what I have learned from you, and I truly hope we will celebrate many scientific breakthroughs together. More than anything, I hope you will soon receive the scientific recognition you so fully deserve.

Marco Balducci
Turku, May 2025

Table of Contents

Acknowledgments	8
Table of Contents	10
List of Original Publications	12
1 Introduction	13
2 Sex Differences in Mathematics, Reading, and Science	17
2.1 Overall Scores	17
2.2 Intraindividual Academic Strengths	18
2.3 Why Study Intraindividual Strengths?.....	19
2.4 Intraindividual Strengths and Ability Tilt.....	21
3 Theories of Sex Differences and Their Changes with Increased Gender Equality	22
3.1 Theories Predicting Reduced Sex Differences with Higher Gender Equality	22
3.2 Theories Predicting Increased Sex Differences with Higher Gender Equality	24
4 Research Design	26
4.1 Study I: Documenting the Gender Equality Paradox.....	26
Data	26
Selection Strategy	27
4.2 Studies II and III: Intraindividual Academic Strengths and the Gender Equality Paradox	27
Data	28
Study II	28
Study III	28
Measures	28
Achievement Scores in PISA.....	28
Gender Equality	29
Wealth	29
Intraindividual Strengths.....	30
Sex Differences in Intraindividual Strengths	31
Study II	31
Study III	31
Analytical Strategy	31

	Study II	31
	Study III	32
5	Results	34
5.1	Study I: Documenting the Gender Equality Paradox.....	34
	Sex Differences in Overall Scores and Associations with	
	Gender Equality Indicators	35
	Mathematics	35
	Reading, Science, and Aggregate Levels	37
	Specific Indicators of Gender Equality	38
	Sex Differences in Anxiety and Attitudes	39
	Sex Differences in Intraindividual Strengths and Gender	
	Equality	40
	The Gender Equality Paradox in Personality Traits	41
	The Big Five and HEXACO Models	41
	Basic Human Values and Vocational Interests	43
5.2	Intraindividual Academic Strengths and the Gender	
	Equality Paradox	44
	Study II.....	44
	Sex Differences in Overall Scores	44
	Sex Differences in Intraindividual Strengths	45
	The Gender Equality Paradox	46
	Study III.....	49
	Sex Differences across PISA Achievement Levels	49
	Intraindividual Strengths	49
	Relative Ratio	51
	Percentages	52
	Correlations among Sex Differences in	
	Intraindividual Strengths	53
	The Gender Equality Paradox.....	55
6	Discussion.....	58
6.1	Summary of the Findings.....	58
	A Review of Sex Differences and Their Relation to Gender	
	Equality	58
	Sex Differences in Intraindividual Strengths across Gender	
	Equality Scores, Time, and Achievement Levels	59
6.2	Implications	60
	Intraindividual Strengths.....	60
	The Gender Equality Paradox	61
6.3	Possible Explanations of the Gender Equality Paradox	62
6.4	Limitations.....	64
6.5	Future Research	65
6.6	Conclusions.....	66
	List of References.....	68
	Original Publications.....	75

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

- I Balducci, M. (2023). Linking Gender Differences with Gender Equality: A Systematic-Narrative Literature Review of Basic Skills and Personality. *Frontiers in Psychology*. 14:1105234.
Doi: 10.3389/fpsyg.2023.1105234
- II Balducci, M., Larose, M.-P., Stoet, G., & Geary, D. C. (2024). The Gender-Equality Paradox in Intraindividual Academic Strengths: A Cross-Temporal Analysis. *Psychological Science*, 35(11), 1246-1259.
Doi: 10.1177/09567976241271330
- III Balducci, M., Larose, M.-P., Stoet, G., & Geary, D. C. The STEM Conundrum: Sex Differences in Intraindividual Academic Strengths and the Gender Equality Paradox Across Academic Achievement Levels. (*under review*)

The original publications have been reproduced with the permission of the copyright holders.

1 Introduction

In recent years, a heated debate has emerged around the topic of how the magnitude of sex differences may change with increased gender equality (i.e., the equal representation of men and women across domains). Some scholars predict these differences to reduce, as they are believed to result primarily from socialization processes that should weaken in more gender-equal countries (A. H. Eagly & Mitchell, 2004; Else-Quest & Grabe, 2012). Conversely, others suggest that men and women in resource-rich, gender-equal environments can more freely express their evolved, intrinsic characteristics (Schmitt et al., 2008). As a result, the gender equality paradox would emerge whereby sex differences would become more pronounced.

The debate has largely centered around educational attributes related to the fields of science, technology, engineering, and mathematics (STEM), owing to their association with sex disparities in long-term labor market outcomes, such as the gender pay gap (Cech, 2013). Despite significant progress toward gender equality in many Western countries, men and women continue to follow sex-specific educational paths. Men are overrepresented in potentially high-paying, high-status, inorganic STEM fields (e.g., computer science), while women tend to prefer the humanities and social sciences, which generally offer lower income prospects (Barone & Assirelli, 2020; van de Werfhorst, 2017).

Many STEM-related educational attributes have been proposed as key contributors to the underrepresentation of women in STEM fields. These encompass sex stereotypes; family influence; sex differences in personality; and disparities in academic skills, such as overall scores in mathematics, reading, and science (Geary et al., 2019; Moè et al., 2021; Wai et al., 2018). Recently, sex differences in intraindividual academic strengths have also been proposed as a contributor due to their close association with educational choices (Dekhtyar et al., 2018). They refer to students' relative advantage in mathematics, reading, or science compared with their overall academic achievement (i.e., mean across these three subjects) and independent of peers' scores (Stoet & Geary, 2018).

Research has shown that girls generally perform better in reading compared with mathematics and science. In other words, reading is their main intraindividual

strength. Meanwhile, boys perform better in mathematics or science compared with reading (Dekhtyar et al., 2018; Stoet & Geary, 2020). Using data from the 2015 wave of the *Programme for International Student Assessment (PISA)*, Stoet and Geary (2018) found that girls' advantage in reading as an intraindividual strength was universal (i.e., observed across countries) and that boys' advantages in mathematics and science were nearly universal. Stoet and Geary also observed that sex differences in reading and science as intraindividual strengths follow a pattern in line with the gender equality paradox; that is, they increase in more gender-equal countries. Conversely, differences in mathematics emerged irrespective of country-level gender equality.

According to *expectancy-value theory* (Eccles, 1983), intraindividual strengths should contribute to educational choices. Individuals with relatively better skills in mathematics and science compared with reading would more likely enroll in a STEM degree. Meanwhile, individuals with relatively better skills in reading compared with mathematics and science would more likely enroll in a humanities or social sciences degree (Bernstein et al., 2019). Indeed, findings from research focused on mathematically precocious youth (Webb et al., 2002), as well as several large-scale studies (e.g., Dekhtyar et al., 2018; Humphreys et al., 1993), support this prediction.

However, despite their significance, intraindividual strengths remain a largely understudied topic. For example, it is unclear whether these strengths follow the same patterns observed for similar measures (e.g., overall scores in mathematics, reading, and science) or whether sex differences persist consistently across time and levels of academic performance. Some scholars have also questioned the replicability of the gender equality paradox in this domain and called for a more in-depth study (Richardson et al., 2020). Richardson and colleagues (2020) argue that to ensure the paradox is not a statistical artifact, a cross-temporal analysis should be the preferred approach.

This dissertation aims to shed light on sex differences in intraindividual academic strengths and the related gender equality paradox to advance the debate on the underrepresentation of women in STEM fields. Accordingly, the following study objectives were pursued: first, the aim was to conduct a comprehensive review of the gender equality paradox in STEM-related domains by examining the scientific literature linking sex differences in academic skills (mathematics, reading, and science) and personality with gender equality indicators (Objective 1). Then, building upon Stoet and Geary's (2018) research, this study aimed to replicate and extend the findings on sex differences in intraindividual strengths and the gender equality paradox by analyzing them over time (Objective 2). Finally, I investigated sex differences in intraindividual strengths and the gender equality paradox at various levels of academic achievement (Objective 3). Note that research on related measures—such as overall scores and ability tilt (numerical-minus-verbal skills, e.g.,

mathematics – reading)—has consistently shown that sex differences tend to be more pronounced among high achievers (Halpern et al., 2007; Wai et al., 2018; Zhou et al., 2017). This pattern, if observed also for intraindividual strengths, could be particularly relevant for sex disparities in STEM fields due to the cognitive demands in these fields.

For Objective 1, I employed a systematic narrative approach based on the PRISMA guidelines to identify, select, and present the relevant scientific literature. For Objectives 2 and 3, using Spearman correlations and linear regression (ordinary least squares (OLS)), I analyzed data from five recent waves (2006–2018) of the PISA, which included over 80 countries and 2.5 million adolescents. To test for the gender equality paradox, sex differences in intraindividual strengths were examined based on the Global Gender Gap Index (GGGI) from the World Economic Forum. In addition, country-fixed and random effects were also included in the models to capture changes over time. Lastly, Stoet & Geary (2018) proposed that many women in developing nations pursue STEM careers due to financial reasons, a pattern less evident in more wealthy, gender-equal countries. Thus, I also considered the contribution of country wealth, calculated by PISA, to the gender-equality paradox in addition to the GGGI.

A literature review revealed that the gender equality paradox has been extensively analyzed across various themes, including overall performance in mathematics, reading, and science, as well as many personality models. The review also highlighted the lack of studies on sex differences in intraindividual strengths and their potential key role in partially explaining the underrepresentation of women in STEM fields, especially in more gender-equal countries.

Analyses of PISA data showed that girls' intraindividual strength in reading and boys' strength in mathematics are universal, while boys' strength in science is nearly universal. Accordingly, across all countries and PISA waves, sex differences in reading and mathematics as intraindividual strengths always favor girls and boys, respectively. However, only a few exceptions have been found for boys' relative advantage in science.

When examining the patterns across different PISA achievement levels (top 5%, bottom 5%, and average), sex differences closely resembled those observed for mean intraindividual scores, with minimal variation. At all achievement levels, no country showed boys' advantage in reading as an intraindividual strength, while only in a few countries, girls were found to outperform boys in mathematics and science.

Additionally, sex differences in reading and science as intraindividual strengths, both overall and across achievement levels, were found to be more pronounced in countries with higher GGGI, aligning with previous research on the gender equality paradox. In contrast, the relationship between sex differences in mathematics and

gender equality never reached statistical significance, suggesting that these differences persist regardless of a country's level of gender equality.

The discussion in this dissertation centers on the results and their theoretical and practical implications, comparing social constructivist and evolutionary theories in relation to intraindividual strengths and the gender equality paradox. Ultimately, it highlights the limitations of previous research and proposes new policy approaches aimed at addressing the underrepresentation of women in STEM fields.

Building on this reflection, I suggest a new framework that incorporates insights from recent research on intraindividual strengths, sex differences, as well as related socio-cultural and evolutionary influences. This framework will not only offer a more nuanced understanding as to why women might be still underrepresented in STEM but also serve as a practical guide for policymakers seeking more effective solutions.

Of note, throughout the dissertation, I use the term "sex" rather than "gender." This choice is intentional, as the focus here is on the biological constituents of sex. In fact, while gender encompasses a broad range of social and cultural identities, sex specifically refers to the physiological distinctions between men and women that are often referred to in this line of research. Accordingly, sex differences are defined as the algebraic difference between the means of women and men on a national level.

2 Sex Differences in Mathematics, Reading, and Science

The underrepresentation of women in STEM has been the subject of a long-standing debate, with explanations ranging from sex stereotypes (Else-Quest & Grabe, 2012), to differences in preferences (Su et al., 2009) to STEM-related cognitive abilities (Halpern, 2013). The discussion around cognitive abilities has often focused on sex differences in mathematics, reading, and science skills, as these are key determinants of academic achievement and long-term educational choices. (Eccles & Wang, 2016; Wai, Lubinski, et al., 2010). In fact, mathematics and science might be linked to careers in STEM fields, while reading might be related to opting out of these fields and pursuing the humanities and social sciences instead.

2.1 Overall Scores

Overall scores generally refer to a composite measure that aims to represent a student's proficiency in mathematics, reading, or science. When analyzing these scores, sex differences are typically observed in the mean and/or performance distribution.

While differences in mean mathematics achievement approach zero, men's advantage in solving more complex mathematical problems is generally observed (Reilly et al., 2019). Additionally, men display greater variability, meaning there is a higher proportion of them at both the highest and lowest ends of the performance distribution, while women are more centered around the mean (Halpern et al., 2007; Wai, Cacchio, et al., 2010). Similarly, men slightly outperform women in science achievement and also show more variability (Halpern, 2013). This higher male variability is not unique to overall scores in mathematics and science but is widely observed across domains (Halsey et al., 2024). It is also significant vis-à-vis the underrepresentation of women in STEM fields. In fact, individuals scoring at the top of the performance distribution in mathematics or science are more likely to pursue STEM fields (Wai, Lubinski, et al., 2010; Zhou et al., 2017), though other factors also play a role (Kahn & Ginther, 2017). Put simply, while average performance differences between men and women are small, the overrepresentation of men

among top performers may partially explain sex disparities in STEM participation (Halpern et al., 2007).

Attitudes, test anxiety, and self-efficacy also contribute to mathematics and science performance (Ashcraft & Moore, 2009; Geary et al., 2019; Esposito et al., 2025). Thus, sex differences in these domains can have a significant influence on shaping educational trajectories (Moakler & Kim, 2014; Sax et al., 2015). Overall, sex differences favor men who, compared with women, are consistently found to experience greater enjoyment and have more positive attitudes toward mathematics and science (Devine et al., 2012; Ganley & Vasilyeva, 2011).

In contrast, women outperform men in verbal tasks—an advantage that might steer them away from STEM fields (Guiso et al., 2008). This pattern is not only larger in magnitude compared with men’s advantage in mathematics and science but is also more reliably observed (Moè et al., 2021). For instance, research on language acquisition shows that by the age of two, girls tend to have a significantly broader vocabulary than boys (Halpern, 2013; van der Slik et al., 2015). Sex differences in verbal tasks are especially pronounced in reading skills. Adolescent girls perform considerably better than boys in reading tests, with the magnitude of this difference estimated to be three times larger than boys’ advantage in mathematics (Stoet & Geary, 2013).

Mathematics, reading, and science are not independent of one another. Higher mathematics proficiency is associated with stronger reading skills and *vice versa*, indicating an interrelation between these cognitive abilities (Bos et al., 2012; Reilly et al., 2019). When examining sex differences in the aggregate level of mathematics, reading, and science, women, overall, tend to achieve higher scores than men, although the latter once again exhibit higher variability (Bergold et al., 2017; Halpern et al., 2007).

2.2 Intraindividual Academic Strengths

Even though a student’s mean academic abilities across mathematics, reading, and science may be categorized as low in relation to peers, performance in one subject can still be relatively stronger compared with the other two subjects. In other words, intraindividual strengths refer to students’ relative advantage in mathematics, reading, and science independent of their mean performance and how they fare in comparison to peers (Stoet & Geary, 2018). For example, a student might have a mean academic achievement level that falls below the national average when considering mathematics, science, and reading scores together. Nonetheless, the student performs relatively better in mathematics compared with reading and science, respectively. In this case, the student’s strongest intraindividual strength is in mathematics, followed by reading, with the least strong intraindividual strength in

science. Specifically, intraindividual strengths represent deviations from the expected performance in each subject based on the aggregate level of achievement (see 2.4 for more details).

Regarding sex differences, research suggests that boys tend to have higher intraindividual strengths in mathematics and science, while girls tend to have higher strength in reading (Dekhtyar et al., 2018; Fellman et al., 2021). Put differently, boys tend to perform better in mathematics and science relative to their own aggregate level of academic abilities, while girls tend to perform better in reading. These patterns seem to be consistent across countries and appear to be universal for reading and nearly universal for mathematics and science (Stoet & Geary, 2018).

2.3 Why Study Intraindividual Strengths?

According to Stoet and Geary (2018), more women should be pursuing STEM degrees, considering their overall mathematics performance, also suggesting that intraindividual strengths may partially explain the underrepresentation of women in STEM participation. Further, other scholars have shown that high school performance correlates with later success, but the specific area of success often aligns with whether the student's strength lies in mathematics or reading (Park et al., 2007). Those with stronger skills in mathematics compared with reading are more likely to succeed in STEM, while those with stronger skills in reading compared with mathematics are more likely to succeed in the humanities. These findings indicate that, although relevant, examining sex differences in overall academic performance alone may be insufficient to fully account for the variation attributed to cognitive, academic domains, potentially contributing to the underrepresentation of women in STEM. Instead, the variation may be better understood by considering intraindividual strengths.

In line with a possible substantial relation to the underrepresentation of women in STEM fields, *expectancy-value theory* (Eccles, 1983) suggests that students' intraindividual strengths, among other factors, are important determinants of academic choices. Notably, expectancy-value theory recognizes the role of cultural influences in shaping students' career decisions (Wang & Degol, 2013). However, this study focuses on the proximate aspects of the theory. In this context, *expectancy* refers to a student's belief about how well they can perform in subjects such as mathematics, reading, or science. Conversely, *value* captures the perceived importance or long-term benefit of engaging in one of the three subjects; this includes both intrinsic enjoyment and personal significance. Thus, students with an intraindividual strength in mathematics or science should be more likely to pursue STEM fields because they expect to perform well and find personal value in them. Meanwhile, students with an intraindividual strength in reading would be more likely

to pursue humanities or social sciences, motivated by similar expectations (Bernstein et al., 2019). In sum, intraindividual strengths can shape academic paths by influencing both the confidence to succeed and the perceived worth of a field. It follows that more men than women would be expected to pursue STEM degrees based on the sex differences in intraindividual strengths reported above.

Although some have questioned the relevance of relative advantages such as intraindividual strengths in explaining sex disparities in STEM fields (Riegle-Crumb et al., 2012; Stearns et al., 2020), the majority of evidence appears to support the expectancy-value proposition. For instance, in their analysis, Dekhtyar et al. (2018) examined data from 167,776 individuals from Sweden and concluded that sex differences in intraindividual strengths correlated with sex disparities in STEM participation, consistent with earlier findings by Jonsson (1999) on the same population. Similar results were reported by Stoet and Geary (2018) in their analysis across 67 countries based on PISA 2015 data ($N = 472,242$). More recent research suggests that sex differences in relative strengths might explain up to 50% or even 80% of the sex disparities in intensive STEM fields (Breda & Napp, 2019; Uunk, 2024). However, doubts have been raised about the real magnitude of this effect (Wan et al., 2023). Further, findings from the Study of Mathematically Precocious Youth (Webb et al., 2002), as well as large-scale studies in the United States (Humphreys et al., 1993), Greece (Goulas et al., 2024), Australia (Justman & Méndez, 2018), and Denmark (Hanushek et al., 2023), also align with expectancy-value theory, providing additional evidence that intraindividual strengths may be more relevant than overall scores to explain the underrepresentation of women in STEM.

Further emphasizing the importance of intraindividual strengths, research shows that sex differences in these domains are nearly twice as large as those observed in overall scores (Stoet & Geary, 2018). Accordingly, Stoet and Geary (2020) found that in 69% of cases, students' sex could be accurately predicted based on their patterns of intraindividual strengths. That is, by examining how students perform in mathematics, reading, and science in relation to their mean academic achievement, it is possible to predict whether the student is a boy or a girl with considerable accuracy. Further, Fellman et al. (2021) demonstrated that growing up in predominantly "male" or "female" environments, that is, having same- or opposite-sex siblings, had no significant impact on these strengths. This outcome suggests that intraindividual strengths are not only consistent across countries and more pronounced than overall scores but also possibly resistant to environmental influences. However, these findings should be interpreted with caution, as intraindividual strengths are a novel concept and may still be shaped by family or societal mechanisms that have yet to be identified.

2.4 Intraindividual Strengths and Ability Tilt

Intraindividual strengths are closely related to ability tilt, though they differ both theoretically and mathematically. Ability tilt refers to a student's relative advantage in one domain over another, without considering overall performance. These domains are typically somehow opposing, such as numerical skills versus verbal skills. For example, a mathematical tilt indicates better performance in mathematics (numerical domain) than in reading (verbal domain). Ability tilt is calculated as the difference between the scores in these domains, typically through a subtraction. One implication is that students might score higher in a different domain not considered in the calculation, that is, the best subject might be overlooked in ability tilt computation. Additionally, ability tilt does not fully capture the relative advantage described by expectancy-value theory, as it does not consider overall performance, which is a key element of this theory.

In contrast, intraindividual strengths measure the deviation of each domain from a student's general ability level. In other words, unlike ability tilt that compares two domains, intraindividual strengths also consider the student's overall performance. These strengths are calculated as the difference between the mean performance across all domains and the score in each specific domain. For example, in a test such as PISA, which assesses mathematics, reading, and science skills, a student's intraindividual strength in reading would be their reading score minus the average performance across all three domains. In other words, overall performance represents the expected score in each subject, while intraindividual strengths show how much a student deviates from this expectation. Considering the example above, if the student's main intraindividual strength is in reading, this means they perform better in reading compared with mathematics and science, and not just above expectations in reading. Therefore, intraindividual strengths seem to better align with expectancy-value theory, as they reflect ability tilt while also accounting for overall performance.

Another advantage of intraindividual strengths is that the concept can be extended beyond mathematics, reading, and science to include additional subjects related to academic skills, such as problem-solving and critical thinking. This broader approach would allow intraindividual strengths to account for mean scores across all considered subjects, offering a potentially more comprehensive understanding of individual abilities than ability tilt which is always limited by pairwise comparisons.

3 Theories of Sex Differences and Their Changes with Increased Gender Equality

Some argue that sex differences arise from socially constructed norms and distinct socialization processes imposed on men and women (A. H. Eagly & Wood, 1999; Else-Quest et al., 2010), while others emphasize human evolutionary history (Schmitt et al., 2008; Stewart-Williams & Halsey, 2021). Regardless of their origins, theories differ on whether these differences should decrease or increase as societies become more gender equal. Below, I first examine theories that propose a decline in sex differences with greater gender equality before discussing those that predict the opposite trend.

3.1 Theories Predicting Reduced Sex Differences with Higher Gender Equality

In their influential article “*Doing Gender*,” West and Zimmerman (1987) examine the origins of sex differences by distinguishing between sex, sex category, and gender. While sex is based on biologically agreed criteria, they argue that sex category refers to the "socially required identificatory displays" that signal one's membership in a particular category (West & Zimmerman, 1987). According to this framework, gender involves the behaviors individuals perform to meet societal expectations associated with their sex category. The authors argue that "doing gender" is inevitable because sex categorization serves as the primary mechanism for differentiation. This categorization is used to justify social arrangements based on the assumption that men and women are intrinsically different (West & Zimmerman, 2009). Consequently, any sex difference derived from sex categorization is viewed as natural and inevitable.

Social role theory (A. H. Eagly & Wood, 1999) embraces the abovementioned view and goes even further by theorizing the origins of this sex categorization. It posits that the interaction between evolved physical traits and the sociocultural context in which these traits are expressed generates sex differences. This interaction is reinforced, in turn, by socio-psychological processes. Evolved physical

predispositions shaped the domestic division of labor in hunter-gatherer societies. Men, owing to their greater strength, carried out more physically demanding activities, while women, because of their ability to breastfeed, were primarily responsible for nurturing tasks (A. H. Eagly & Wood, 2012). As societies evolved, this division of labor adapted to the developing social environment. Physically demanding activities progressively transformed into paid, power-related roles, and men were increasingly viewed as more suited for them. Meanwhile, women continued to be associated with nurturing tasks, reinforcing the perception that they were not fit for roles involving authority (Wood & Eagly, 2013).

Social psychological processes strengthen the division of labor by making differences between men and women appear “natural” (A. H. Eagly & Mitchell, 2004). Based on observed behaviors, individuals form expectations and generate specific sex roles. Sex roles are then learned, internalized, and maintained through reward and punishment mechanisms that influence individuals’ choices (Friedman & Downey, 2002). In other words, sex differences are theorized to emerge due to constructed roles imposed on men and women by prevailing social structures.

The *gender stratification hypothesis* (Baker & Jones, 1993) aligns with social role theory. It posits that essentialist beliefs about men and women interact with social structures to produce sex differences. Specifically, individuals are more likely to engage in tasks they value and believe they can succeed in. However, essentialist beliefs shape these two factors, because of which some tasks are seen as more appropriate for either sex. Consequently, women may feel less confident about their potential for success in, and even possibly avoid, traditionally masculine fields, such as STEM. Men may show a similar tendency toward fields perceived as feminine, such as education or nursing. Accordingly, essentialist beliefs are maintained, perpetuating sex differences (Frome & Eccles, 1998). Essentialist beliefs are further reinforced by families, peers, and the media, where men and women are expected to conform to sex-specific roles (Bussey & Bandura, 1999).

According to the abovementioned theories, sex differences are largely the result of stereotypes and societal expectations. When they diminish, as is the case in more gender-equal societies, differences between men and women are predicted to decrease (A. H. Eagly & Mitchell, 2004; Else-Quest et al., 2010). Indeed, evidence supporting a reduction in sex differences in more gender-equal countries has recently emerged, especially concerning partner preferences, sexual behaviors, and segregation in the labor market (Almstedt Valldor & Halldén, 2023; Herlitz et al., 2024).

3.2 Theories Predicting Increased Sex Differences with Higher Gender Equality

In contrast to the theories outlined above, some scholars have argued that sex differences may instead widen as gender equality increases. *Gender essentialism* (Charles & Bradley, 2009) suggests that individualism and self-expression in gender-equal countries encourage personal preferences but the origins of these preferences remain unquestioned. In this vein, preferences are believed to be shaped by essentialist beliefs about men and women that reinforce pre-existing, socially imposed sex roles and influence individuals accordingly. As sex roles are reinforced, sex differences are also perpetuated and even amplified (Levanon & Grusky, 2018).

Further, societal systems, although designed to accommodate “expressive choices,” inadvertently support stereotypes as individuals enact their internalized sex identities instead of subjective preferences (Charles et al., 2014; Rawlings, 2007). For instance, extensive family policies and long-care leaves, typically seen in gender-equal countries, tend to steer women into traditionally feminine roles (Freiberg, 2019). In line with the gender essentialist perspective, some scholars have argued that stronger stereotypes about men and women are found in more gender-equal nations (Breda et al., 2020; Napp and Breda, 2022). Moreover, others suggest that sex differences can be predicted by country-level “cultural individualism” (Bleidorn et al., 2016; Kaiser, 2019).

Further, *evolutionary theories* (Geary, 2010; Schmitt et al., 2017) propose that sex differences would become more pronounced in more gender-equal societies, although for reasons distinct from those proposed by gender essentialists. In these societies, abundant resources and social norms that emphasize self-expression allow individuals to make career choices independent of long-term financial returns. Interestingly, this context would enable individuals to express more freely evolved, sex-typical behaviors, the underlying mechanisms of which are explored below.

A key aspect of evolutionary theories is sexual selection and the related evolution of vulnerability. Certain traits that increased reproductive success were naturally selected and passed down through generations. However, the traits influencing reproductive success differed between men and women due to intersexual choices. In other words, differences in reproductive advantages and mate preferences have shaped sex differences throughout human evolutionary history. Sexually selected traits are often more energetically demanding for the sex that benefits, making them more vulnerable and susceptible to environmental changes. In addition, cultural norms that suppress self-expression may also restrain evolved sex differences. Therefore, in environments with high ecological and cultural constraints, sex differences should be attenuated. Conversely, better economic conditions and value attested to self-expression, generally found in more gender-equal countries, reduce constraints, increase energy availability, and allow sexually selected traits to be fully

expressed (Geary, 2015). Evidence supporting evolutionary predictions has emerged, for example, with physical traits such as height, as well as psychological characteristics such as personality (Costa et al., 2001; Sohn, 2015). It is important to note that evolutionary predictions apply specifically to traits shaped by sexual selection. Other sex differences may decrease or remain stable as gender equality increases. In other words, the theories presented above might not be mutually exclusive.

4 Research Design

This research explored the gender equality paradox in STEM education by examining a specific related aspect: intraindividual academic strengths. In Study I, a theoretical framework, grounded in the current literature, was developed; and Studies II and III involved an in-depth examination of sex differences in these strengths. Specifically, this project addressed the following research questions:

- 1) Do sex differences increase or decrease with higher national gender equality? (Study I)
- 2) Is there a gender equality paradox in intraindividual academic strengths? (Study II)
- 3) Does the strength of the gender equality paradox in intraindividual strengths vary at different levels of academic achievement? (Study III)

4.1 Study I: Documenting the Gender Equality Paradox

Data

To offer a broad perspective on the phenomenon, Study I involved a comprehensive literature review of the gender equality paradox across aspects and domains related to STEM participation. These include not only mathematics, reading, and science but also personality traits that have been consistently linked to long-term career choices (Kuhn & Wolter, 2022; Su et al., 2009). The main goal was to develop a theoretical framework that would support this doctoral project and also serve as a foundation for future research, given that, at the time Study I was initiated, a comprehensive analysis of the gender equality paradox was still lacking. Published scientific articles were accessed and selected from databases such as Scopus, Web of Science, the Social Science Database, and Google Scholar, and chosen based on their focus on sex differences and their relationship to composite and/or specific

indices of gender equality. A total of 31 peer-reviewed papers were identified for this study.

Selection Strategy

To select relevant studies, I followed a systematic approach based on the PRISMA guidelines. The search focused on keywords such as “gender/sex differences in mathematics/reading/science,” “gender/sex differences in personality,” “gender/sex differences in basic human values,” and “gender/sex differences in vocational interests.” The search was further refined by incorporating parameters such as “gender equality,” “egalitarianism,” and “inequality.”

Eligible studies were identified based on the use of quantitative cross-national data to analyze sex differences in relation to composite and/or specific measures of gender equality. Since a novel wave of egalitarian theories known as “new egalitarianism” (Martin, 2017) emerged after the global crisis back in 2008, the time range considered for the research was 2009–2020. Studies were classified based on two main groups: academic skills and personality. Each group was further divided into specific subgroups. For academic skills, the subgroups included mathematics, science (encompassing attitudes and anxiety), reading, and overall scores. For personality, the subcategories included the Big Five traits, the HEXACO model, basic human values, and vocational interests. The final search was conducted in November 2022.

Articles were screened by title, abstract, and keywords. Only articles in English were considered. The selection process was managed under the supervision of two scholars who ensured consistent application of criteria.

Of the 91 articles initially preselected, 35 were excluded after screening due to mismatches with selection criteria, and another 25 were excluded for reporting research on unrelated sex differences. Following exclusions, 31 articles were finally selected.

4.2 Studies II and III: Intraindividual Academic Strengths and the Gender Equality Paradox

Results from the literature review highlighted that while the gender equality paradox has been extensively studied with regard to overall scores for mathematics, reading, and, science, as well as for personality traits, intraindividual academic strengths emerged as a novel and potentially significant area of research, especially concerning STEM participation. Consequently, Studies II and III focused on exploring sex differences in intraindividual academic strengths in greater depth—examining these differences in mean values, at various levels of achievement, and across time.

Specifically, Study II analyzed mean sex differences in intraindividual strengths and their relationship to gender equality, while Study III extended this analysis to different levels of academic achievement. Both studies included multiple PISA waves to assess the consistency of the results over time.

Data

Study II

Study II drew data from five waves of the PISA: 2006, 2009, 2012, 2015, and 2018. Although PISA data have been collected since 2000, only waves for which the GGGI data from the World Economic Forum (more below) were also available were included. The number of participating countries varied across waves, from 57 in 2006 to 77 in 2018. The final sample comprised 2,470,874 adolescents from 75 countries, along with 10 additional municipalities and regions where whole-country data were unavailable. These regions were Indian Himachal Pradesh and Tamil Nadu; Miranda in Venezuela (PISA 2009); Buenos Aires (PISA 2015); and Hong Kong, Macao, and Chinese Taipei (in all five PISA waves).

Study III

Study III also utilized data from PISA, but only the 2012, 2015, and 2018 waves were considered. The 2022 wave was excluded due to cross-country differences in COVID-19 restrictions that could have biased the results. The number of participating countries ranged from 64 in 2012 to 77 in 2018. The final sample included 1,611,512 adolescents from 74 countries and eight regions (Buenos Aires in PISA 2015 as well as Hong Kong, Macao, and Chinese Taipei in all three PISA waves).

Measures

Achievement Scores in PISA

PISA is an international assessment that evaluates adolescents' abilities in mathematics, reading comprehension, and science literacy every three years. Its primary goal is to assess general competencies at the country level across these domains rather than individual performance. Student competencies are drawn from a set of plausible values (PVs), which are estimates generated from the posterior distribution of actual test scores. These PVs are scaled such that the mean score for

member countries of the Organisation for Economic Co-operation and Development (OECD) is set to 500, with a standard deviation (SD) of 100 points.

From 2006 to 2018, PISA made several changes in its assessment. Until 2012, PISA provided five PVs for each domain, which increased to 10 PVs from 2015 onward. Additionally, in 2015, PISA transitioned from a paper-based to a computer-based format, which prevents students from returning to questions they have skipped or answered. Lastly, in 2018, PISA introduced a new two-stage method for assessing reading comprehension: students first completed a basic reading test, and whether they would receive a subsequent harder or easier test was determined based on their initial performance.

Gender Equality

The World Economic Forum has published GGGI data annually since 2006. The GGGI assesses countries' progress toward gender equality in four key domains—economic participation and opportunity, educational attainment, health and survival, and political empowerment—using 14 indicators. For each indicator, scores are capped at 1. Scores below 1 indicate an advantage for men, while a score of 1 represents sex parity or women's advantage. The GGGI covers over 140 countries, making it the most comprehensive gender equality indicator for cross-country comparisons.

While various gender equality indicators exist, the GGGI offers extensive country coverage and a comprehensive approach to gender equality. Its strength lies in incorporating diverse indicators, offering a nuanced understanding of country-level factors potentially linked to individuals' freedom of choice. In the PISA country sample for Study II, GGGI scores ranged from 0.583 for Turkey in 2006 to 0.881 for Iceland in 2015. In Study III, scores varied from 0.593 for Jordan in PISA 2012 to 0.881 for Iceland in 2015. Notably, GGGI scores demonstrate high stability across PISA waves ($r \geq .85$).

Wealth

The gender equality paradox may stem not from gender equality itself but from the better economic conditions generally found in more gender-equal countries (Stoet & Geary 2018). Gender equality and economic development are indeed closely linked, thus wealth was also factored into the analyses in Study II. PISA does not directly measure economic prosperity at the country level; instead, it collects data on family possessions. In the student questionnaires, participants report whether they have their own room, internet access, and other household items. These items are then combined into an index called “WEALTH,” which reflects family affluence.

Using the PISA family wealth index, a measure of each country's economic prosperity was derived. First, the weighted mean of the index for each country was computed, generating a new variable, "Country Wealth." Then, the new variable was standardized across countries ($M = 0$, $SD = 1$). This measure more directly reflects national-level economic prosperity than traditional metrics such as gross domestic product (GDP). Nevertheless, note that the country wealth variable correlates strongly with GDP per capita ($r \geq .80$) based on purchasing power parity (constant 2017 international dollars) from the *World Economic Forum*.

Intraindividual Strengths

As mentioned above, PISA provides student scores in mathematics, reading, and science as PVs. Five PVs were provided for each domain up until PISA 2012. These increased to 10 PVs from PISA 2015 onward. PVs are used to evaluate each student's best, second-best, and lowest performance in the academic subjects and determine their intraindividual strengths. To do so, I employed the method outlined by Stoet and Geary (2018), summarized below. Following the PISA guidelines (OECD, 2009), every step was first applied separately for each PV, before averaging the results.

- 1) I created new PVs reflecting the average performance across subjects, that is, $(PV1Math + PV1Reading + PV1Science) / 3$. The same procedure was carried out for each of the five or 10 PVs. These average performance scores were then standardized ($M = 0$, $SD = 1$) within each country and wave, creating a new variable, $zGen$.
- 2) Similarly, PVs for each subject (mathematics, reading, and science) were standardized on a country-by-country and wave-by-wave basis, yielding new standardized scores ($zMath$, $zRead$, and $zSci$) with a mean of 0 and an SD of 1.
- 3) Intraindividual strengths were then calculated as the difference between each subject's z-score and $zGen$. For example, $z1Math - z1Gen$ represents the mathematics intraindividual strength for the first PV. The same calculations were carried out for each of the remaining PVs and also for $zRead$ and $zSci$.

Sex Differences in Intraindividual Strengths

Study II

Study II examined the mean sex differences in intraindividual strengths. After Step 3 described above, the average scores in intraindividual strengths in mathematics, reading, and science for boys and girls were calculated for each country. Then, the scores of boys were subtracted from those of girls to compute sex differences in intraindividual strengths at the country level.

Study III

In Study III, I analyzed sex differences in intraindividual strengths at different levels of PISA achievement. Accordingly, first, I computed the average PISA performance across zGen scores (Step 1 above). Then, I divided the sample based on three levels of PISA achievement: high achievers = top 5% (95th percentile and above, inclusive); low achievers = bottom 5% (5th percentile and below, inclusive); and average achievers (> 5th to < 95th percentile).

At each PISA achievement level, sex differences in intraindividual strengths were calculated for every country by subtracting the mean scores of girls from those of boys. Additionally, this procedure was repeated using a 10th percentile cut-off, dividing the sample into the top 90th percentile and bottom 10th percentile.

Analytical Strategy

Study II

The analytical strategy for Study II comprised four steps. I adhered to PISA's guidelines for complex survey data analysis and used an alpha criterion of .05. All analyses were performed in STATA 17.

- 1) **Sex Differences in Overall Scores:** Country-by-country sex differences in overall scores (not intraindividual strengths) in mathematics, reading, and science were calculated using PISA's *repest* command, which employs balanced repeated replication (BRR) weights for unbiased estimation. Computing sex differences in overall scores also provides information as to whether sex differences in intraindividual strengths align with or diverge from them.
- 2) **Sex differences in Intraindividual Strengths by PISA Wave:** Sex differences in intraindividual strengths were analyzed for each PISA wave individually to

include all countries. Some countries, in fact, participated in only one or a few assessments.

- 3) **Exploring the Gender Equality Paradox:** To explore the gender equality paradox, I used Spearman's p correlations (r_s) for associations between sex differences in intraindividual strengths and GGI, as well as wealth, followed by linear regression models (OLS) using a bootstrap method with 1,000 iterations.
- 4) **Tracking the Gender Equality Paradox Over Time:** For this step, the analysis focused solely on countries that participated in all five PISA waves. Longitudinal linear regression models with random (including a GGI-by-wave interaction) and country-fixed effects were used to assess the gender equality paradox over time.

Study III

Study III employed a three-step analytical approach to extend the findings of Study II, using an alpha criterion of .05. The analyses were performed in STATA 18.

- 1) **Sex Differences by Achievement Level:** Mean sex differences, across countries, in mathematics, reading, and science were computed per PISA wave and at each achievement level. For the achievement levels, I computed two sets of cutoffs and ran the analyses for each: top 5%, bottom 5%, and average, as well as top 10% and bottom 10%. An analysis of variance (ANOVA) tested for significant differences across achievement levels. Note that here, achievement levels are not the same as PISA proficiency levels. Proficiency levels assess mathematics, reading, and science skills separately, with different cutoffs for each skill, making direct comparisons of intraindividual strengths difficult. Using PISA's student weights, I also calculated both the girl-to-boy ratio and the percentage of girls and boys having mathematics, reading, or science as the subject linked to their main intraindividual strength.
- 2) **Intraindividual Strength Correlations:** Using Spearman's p correlations (r_s), I tested for associations between intraindividual strengths by correlating sex differences in mathematics as an intraindividual strength with those in reading and science at each PISA achievement level. The same procedure was repeated for sex differences in reading and science as intraindividual strengths.
- 3) **Gender Equality Paradox by Achievement Level:** Lastly, I examined whether the gender equality paradox occurred consistently across achievement levels.

This step mirrored the procedure in Study II, Step 3, but this time, it was applied to each PISA achievement level separately.

5 Results

The first section of the results summarizes the findings from the literature review, covering Part 1 of the thesis. Subsequently, the results from both Studies II and III, constituting Part 2, are presented.

5.1 Study I: Documenting the Gender Equality Paradox

Two main standardized international tests are generally used to assess skills in mathematics, reading, and science: Trends in International Mathematics and Science Study (TIMSS) and PISA. TIMSS assesses the mathematics and science academic achievements of students in the fourth and eighth grades worldwide every four years, while PISA assesses not only mathematics and science but also reading comprehension among adolescents (age range: 15 years, 3 months to 16 years, 2 months) every three years.

Sex differences resulting from the abovementioned tests are then related to indices of gender equality. These indices can be either composite, encompassing multiple indicators, or specific, focusing on a single area. While numerous composite international indices of gender equality exist, the GGGI from the World Economic Forum and the United Nations' Gender Empowerment Measure (GEM) are the most widely used. Meanwhile, specific indicators of gender equality typically evaluate women's economic or political empowerment, as well as their representation in a specific field of interest.

The following sections start with descriptions of analyses focused on the relationship between gender equality indicators and overall scores in mathematics, reading, and science. This includes an analysis of aggregate scores (combining mathematics, reading, and science) as well as attitudes and anxiety levels. They then detail analyses of the gender equality paradox considering personality factors, first with respect to the Big Five and HEXACO models, followed by basic human values and vocational interests.

Sex Differences in Overall Scores and Associations with Gender Equality Indicators

Mathematics

When examining variations in sex differences in mathematics in the TIMSS test, there appears to be no significant correlation with composite indices of gender equality. In other words, sex differences in mathematics on TIMSS assessments emerge regardless of countries' level of gender equality. However, a more complex pattern emerges when considering PISA (Tab. 1).

At first glance, sex differences in mathematics seem to decrease in countries that have made significant strides in gender equality, aligning with social role theory and the gender stratification hypothesis (Else-Quest et al., 2010; Hyde & Mertz, 2009). However, there is variation in the effect sizes reported across studies and the correlation coefficients consistently remain quite small ($r = 0.09\text{--}0.14$).

Some scholars have proposed that this correlation might depend on the year when the data were collected or the specific countries included in the sample. Stoet and Geary (2013, 2015) found that social role theory and the gender stratification hypothesis were supported only by the 2003 PISA assessment. Nevertheless, other PISA waves showed no significant relationship between gender equality composite indicators and sex differences in overall mathematics performance. Others have also demonstrated that including or not including Scandinavian, Muslim-majority, or gender-equal countries, where girls outperform boys, might influence the results, leading to significant fluctuations in the estimates (Fryer & Levitt, 2010; Kane & Mertz, 2012; Stoet & Geary, 2015).

Table 1. Sex differences in mathematics (men > women) and relation to both composite gender equality indices (top) and specific indicators (bottom).

<i>Composite Indices</i>					
Reference	Test	Year	Countries	Indices	Correlation
(Else-Quest et al., 2010)	TIMMS, PISA	2003	46, 41	GEM, GEQ, SIGE, GGI	Not Found Negative
(Fryer & Levitt, 2010)	TIMMS, PISA	2003	47, 41	GGI	Not found Negative
(Kane & Mertz, 2012)	TIMMS, PISA	2007 and 2009	52, 65	GGI, GEI	Positive
(Hyde & Mertz, 2009)	PISA	2003	30	GGI	Negative
(Reilly, 2012)	PISA	2009	34	GGI, RSW	Not found
(Stoet & Geary, 2013)	PISA	From 2000 to 2009	75	GEM, GGI	Not reliable
(Stoet & Geary, 2015)	PISA	From 2000 to 2010	41 to 74	GEM, GGI	Not reliable
(Ireson, 2017)	PISA	2012	65	GGI	Not found
(Gevrek et al., 2018)	PISA	2012	56	GGI	Negative
(Anghel et al., 2019)	PISA	From 2003 to 2015	73	GGI	Not found
(Gevrek et al., 2020)	PISA	2012	56	GGI	Not found
(Ghasemi et al., 2019)	TIMMS	2015	48	GGI	Not found
(Reilly et al., 2019)	TIMMS	2011	45	GGI	Not found
<i>Specific Indicators</i>					
(Else-Quest et al., 2010)	TIMMS, PISA	2003	46, 41	RE, WR, WPEA, FPS, HMP	Not found Negative
(Reilly, 2012)	PISA	2009	34	WR	Negative
(Penner and Cadwallader Olsker, 2012)	TIMMS	1995	22	WE, WL	Negative Positive
(Stoet & Geary, 2015)	PISA	From 2000 to 2010	41 to 74	WR, WPEA, FPS	Negative Positive
(Gevrek et al., 2020)	PISA	2006, 2012	56	RE, WR, WPEA, FPS, HMP	Negative Not found

GEM = Gender Empowerment Measure; GEI = Gender Equality Index; GGI = Gender Gap Index; GEQ = Gender Equality and Quality of Life; SIGE = Standardized Index of Gender Equality; RSW = relative status of women; RE = ratio of men to women in education; WR = women in research; WPEA = women’s participation in economic activities; FPS = female parliamentary seats; HMP = female’s higher labor market positions; WE = women’s parity in education; WL = women’s labor market participation. (Table reproduced from: doi.org/10.3389/fpsyg.2023.1105234).

Reading, Science, and Aggregate Levels

Only a few studies have focused on sex differences in reading and their relation to gender equality composite indicators; however, there seems to be significant agreement among scholars (Tab. 2). The findings consistently indicate that sex differences favor women and increase with improvements in national gender equality, supporting the gender equality paradox (Gevrek et al., 2020; Reilly, 2012).

Conversely, research on sex differences in science achievement has yielded mixed results, supporting neither theories predicting an increase nor a decrease in sex differences with greater gender equality. In an early analysis, Reilly (2012) found a negative relation between sex differences in science achievement and gender equality ($r = -0.29$). However, in line with a higher male variability, he also noted that men were overrepresented among high performers. In contrast, later studies failed to replicate this finding, suggesting that sex differences in science achievement persist regardless of a country's level of gender equality (Ireson, 2017). Yet, more recently, evidence for the gender equality paradox in science has emerged—Reilly et al. (2019) reported that men tend to outperform women in more gender-equal societies, while women tend to outperform men in nations with lower gender equality.

Studies examining sex differences in aggregate levels of mathematics, reading, and science skills are currently insufficient to draw any firm conclusion. Furthermore, the existing literature has produced mixed results that do not consistently align with theories predicting either an increase or a decrease in sex differences with greater gender equality. Analyzing the PISA 2003 wave, Stoet & Geary (2015) found that sex differences in aggregate skill levels are positively related to the GGGI, indicating that these differences tend to increase with improvements in gender equality. However, similar to mathematics, the correlation seems to depend on the specific countries sampled and year analyzed. For instance, this correlation significantly weakened or disappeared when either Iceland or Finland was excluded or when considering other PISA waves (Ireson, 2017; Stoet & Geary, 2015).

Table 2. Sex differences in science (men > women), reading, and overall scores (men < women) and relation to gender equality indices (top) and specific indicators (bottom).

Science					
<i>Composite Indices</i>					
Reference	Test	Year	Countries	Indices	Correlation
(Reilly, 2012)	PISA	2009	34	GGI	Negative
(Ireson, 2017)	PISA	2012	65	GGI	Not found
(Reilly et al., 2019)	TIMMS	2011	45	GGI	Positive
<i>Specific Indicators</i>					
(Reilly, 2012)	PISA	2009	34	WR, RSW	Negative/ Not found
Reading					
<i>Composite Indices</i>					
(Reilly, 2012)	PISA	2009	34	GGI	Not found
(Gevrek et al., 2020)	PISA	2006, 2012	56	GGI	Positive
<i>Specific Indicators</i>					
(Reilly, 2012)	PISA	2009	34	WR, RSW	Positive
(Gevrek et al., 2020)	PISA	2006, 2012	56	RE, WR, WPEA, FPS, MP	Positive
Overall Scores					
<i>Composite Indices</i>					
(Stoet & Geary, 2015)	PISA	From 2000 to 2010	41 to 74	GEM, GGI	Not found
(Ireson, 2017)	PISA	2012	65	GGI	Not found
(Eriksson et al., 2020)	PISA, TIMMS	From 2003	74	GGI	Positive
(Stoet & Geary, 2020)	PISA	From 2009 to 2015	55 to 71	GGI	Positive
<i>Specific Indicators</i>					
(Bergold et al., 2017)	TIMMS PIRLS	2011	32	WAE, TSER, WPLM, WR	Positive/ Negative

GEI = Gender Equality Index; GGI = Gender Gap Index; GEQ = Gender Equality and Quality of Life; RSW = relative status of women; WAE = women’s access to education; TSER = tertiary school enrollment rate, boys/girls; WPLM = women’s participation in the labor market; WR = women in research; RE = ratio of men to women in education; WPEA = women’s participation in economic activities; FPS = female parliamentary seats; HMP = women’s higher labor market positions. (Table reproduced from: doi.org/10.3389/fpsyg.2023.1105234)

Specific Indicators of Gender Equality

Some scholars have criticized studies using composite indicators (e.g., GGGI and GEM) to assess how the magnitude of sex differences changes with increases in national gender equality. They argue that these indicators are unreliable metrics as they do not directly measure the most significant subdomains (e.g., the ratio of women to men holding parliamentary seats) of gender equality influencing career choices. They

instead propose using specific indicators, such as the share of executive positions, which they argue more effectively capture the indirect influences on the decision-making processes of men and women (Else-Quest et al., 2010).

Regarding sex differences in mathematics, the findings related to specific indicators appear to align with those for composite indices, supporting neither theories predicting an increase nor a decrease in sex differences with greater gender equality. At first glance, factors such as the number of women in research, female participation in economic activities, the ratio of women to men holding parliamentary seats, and greater educational equality seem to correlate negatively with sex differences in mathematics performance. Specifically, higher levels of these indicators are associated with smaller sex differences in mathematics (Else-Quest et al., 2010; Penner & CadwalladerOlsker, 2012).

However, even findings for specific indicators appear to depend on the year and country sample. Results related to the "women in research" index fluctuate significantly according to the PISA wave analyzed ($r = -0.16$, $r = -0.68$; Reilly, 2012; Stoet & Geary, 2015). Further, countries such as Latvia, Serbia, Tunisia, and Thailand, which are less gender-equal and display lower sex differences, along with non-OECD nations, seem to be driving the correlation (Reilly, 2012; Stoet & Geary, 2015).

For the "women's economic activity" index too, the results appear unreliable and largely dependent on the specific PISA wave analyzed. Higher levels of women's economic activity were initially associated with a reduction in sex differences in mathematics in the 2000 and 2003 PISA waves (Stoet & Geary, 2015). Nevertheless, these results were not replicated in subsequent PISA waves (2006 and 2009). Stoet and Geary (2015) also examined the number of "females in parliamentary seats," but this indicator was never linked to sex differences in mathematics. However, a positive correlation was observed in the PISA 2003 wave when non-OECD or Nordic countries were excluded from the sample (Stoet & Geary, 2015).

In contrast to mathematics, there has been less focus on reading and science when examining the link between sex differences and specific indicators. Research has yet to explore how these indicators relate to sex differences in reading, and only a limited number of studies have addressed sex differences in science. Overall, specific indicators, as composite indices, also yield mixed results regarding sex differences in science. There appears to be no relationship between these differences and the "relative status of women," while a positive relationship has been found for the number of "women in research" ($r = 0.39$; Reilly, 2012).

Sex Differences in Anxiety and Attitudes

As outlined above, overall performance in mathematics, reading, and science can be influenced by anxiety and general attitudes toward academic testing (Devine et al.,

2012). The number of studies that have focused on anxiety and attitudes in mathematics have been considerably higher compared with those examining these variables in reading and science.

When examining sex differences in mathematical attitudes and anxiety and their connection to composite indicators, the findings reveal larger sex differences in more gender-equal countries (Else-Quest et al., 2010; Stoet et al., 2016). Considering the specific indicators, results fluctuate significantly depending on which indicator is analyzed. For instance, an increase in women's political representation is associated with wider sex differences in self-efficacy and motivation in mathematics. Conversely, indicators such as "equality in wages," "parity in secondary and tertiary education," and "equal political representation" predict smaller sex differences (Else-Quest et al., 2010; Gevrek et al., 2020).

To my knowledge, no studies have specifically examined sex differences in reading anxiety and attitudes. Research on science, however, aligns with the gender equality paradox, indicating that sex differences in self-efficacy, enjoyment, and interest in science are larger in more gender-equal nations (Liou et al., 2022; Stoet & Geary, 2018).

Sex Differences in Intraindividual Strengths and Gender Equality

The studies discussed in the previous sections relied on inter-group comparisons—contrasting mean scores of men to those of women—which may not be the most suitable approach for assessing the relationship between sex differences in mathematics, reading, and science and gender equality. In fact, these sex differences are generally small to medium in magnitude, and any specific patterns may not be immediately apparent.

Some scholars have suggested shifting the focus to intraindividual academic strengths, as this approach could advance the discussion on the underrepresentation of women in STEM fields. Intraindividual strengths appear to be larger than those found in overall scores, more stable across countries, and potentially more closely related to STEM careers (Dekhtyar et al., 2018; Stoet & Geary, 2018; Wang & Degol, 2017).

Only one study has focused on intraindividual academic strengths and their relation to gender equality. Stoet and Geary (2018) analyzed data from 472,242 adolescents across 67 countries and regions using the 2015 PISA wave. They examined the relationship between sex differences in mathematics, science, and reading as intraindividual strengths and the GGGI. Their results revealed that while sex differences in mathematics as an intraindividual strength were not related to country-level gender equality and emerged regardless of it, sex differences in reading and science increased with higher GGGI scores.

There has been heated debate about the reliability of the aforementioned findings. Some scholars have criticized the results of Stoet and Geary's study, arguing that they may not be replicable in other PISA waves (Richardson et al., 2020). However, no other study has examined whether the gender equality paradox in these strengths persists over time. Additionally, while intraindividual strengths may be more relevant than overall scores for understanding STEM field participation, few studies have explored sex differences in this dimension in depth.

The Gender Equality Paradox in Personality Traits

Several models of personality are commonly used to evaluate associations with long-term choices and individual characteristics, including the Big Five (openness, conscientiousness, extraversion, agreeableness, and neuroticism) and the HEXACO (honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness) models. Other relevant assessments include basic human values (Schwartz, 2007) and vocational interests (Holland, 1997), which more specifically describe how personality interacts with career choices.

Sex differences have been reported across all the abovementioned personality models, with women generally scoring higher in neuroticism and agreeableness from the BIG five as well as in emotionality and honesty-humility from the HEXACO model (Lee & Ashton, 2020; Murphy et al., 2021). Regarding value priorities and vocational interests, men tend to prioritize power, stimulation, hedonism, achievement, and self-direction, as well as preferring to work with “things” (i.e., inorganic fields), while women prioritize universalism, benevolence, and working with people (Lippa, 2010b; Su et al., 2009).

The relevance of these sex differences lies in their association with career paths. For instance, men, with their tendency toward realistic and investigative interests, are more likely to prefer STEM fields. Meanwhile, women, with their inclination toward social and artistic interests, are more likely to prefer the social sciences and humanities (Kuhn & Wolter, 2022).

The Big Five and HEXACO Models

The gender equality paradox in personality traits has received significant research attention (Tab. 3). Early studies in the 2000s (Costa et al., 2001; Schmitt et al., 2008) first suggested a counterintuitive trend: as gender equality increased, sex differences in neuroticism and agreeableness, from the five-factor model, also increased. Subsequent research has replicated this trend, showing that the paradox was mainly driven by men's lower scores (Lippa, 2010a). A similar pattern has also been found for emotionality in HEXACO assessments (Lee & Ashton, 2020).

However, personality is multifaceted in nature and focusing on individual traits might result in misleading conclusions (Vianello et al., 2013). In this vein, Del Giudice (2009) suggested that to capture the true magnitude of sex differences in personality, the overall variance should be considered. In other words, while individual sex differences in specific traits may be small when considering the collective pattern of differences across personality traits (i.e., overall variance), the overlap between men's and women's personality profiles is significantly reduced. Considering the overall variance in personality, studies have found a strong positive correlation between sex differences and gender equality measures, supporting the gender equality paradox (Kaiser, 2019; Mac Giolla & Kajonius, 2019). Put simply, sex differences in personality are wider in more gender-equal countries.

Table 3. Sex differences in personality (men–women) and relation to composite indices of gender equality.

<i>Big Five and HEXACO Models</i>					
Reference	Model	Year	Countries	Indices	Correlation
(Lippa, 2010b)	Big Five	2005	53	UN	Positive/Not found
(Mac Giolla & Kajonius, 2019)	Big Five	2000	22	GGI	Positive
(Kaiser, 2019)	Big Five	2001 and 2011	50	GGI	Positive
(Lee & Ashton, 2020)	HEXACO	from 2014 to 2018	48	GGI	Positive/Not found
<i>Basic Human Values</i>					
(Schwartz & Rubel-Lifschitz, 2009)		2002 and 2004	25	PCC/ PA	Positive
(Fors Connolly et al., 2020)		2016	32	GEI	Positive
<i>Vocational Interests</i>					
(Ott-Holland et al., 2013)		2013	20	GLOBE Study	Positive/Not found
(Tao et al., 2022)		2021	42	GEI	Positive

GEM = Gender Empowerment Measure; GEI = Gender Equality Index; GGI = Gender Gap Index; GEQ = Gender Equality and Quality of Life; SIGE = Standardized Index of Gender Equality; RSW = relative status of women; RE = ratio of men to women in education; WR = women in research; WPEA = women's participation in economic activities; FPS = female parliamentary seats; HMP = female's higher labor market positions; WE = women's parity in education; WL = women's labor market participation; WAE = women's access to education; TSER = tertiary school enrollment rate, boys/girls; WPLM = women's participation in the labor market; HMP = women's higher labor market positions; UN = UN's Gender-Related Development Index; PCC = Population Crisis Committee Index of Gender Equality; PA = Prescott-Allen's (2001) index of gender equity. (Table reproduced from: doi.org/10.3389/fpsyg.2023.1105234)

Basic Human Values and Vocational Interests

As stated above, men, on average, regard power, achievement, and stimulation more highly compared with women. However, women regard benevolence and universalism more highly compared with men. Past research has shown a trend in line with the gender equality paradox in basic human values, although people of both sexes generally assign less importance to masculine values (Schwartz and Rubel-Lifschitz, 2009). Interestingly, Fors Connolly et al. (2020) showed that although sex differences are wider in more gender-equal countries, there was a convergence between men and women over time in benevolence (with a decrease in sex differences, Cohen's d -15%), while gaps in universalism and stimulation remained constant. This finding highlights the need to incorporate a temporal dimension in further research when analyzing the paradox.

Regarding vocational interests, few studies have examined the association between sex differences and gender equality. For instance, using the Brinkman model of interests, Ott-Holland (2013) and colleagues found that sex differences in musical and persuasive interests were smaller in countries that fared well in gender equality, while differences in clerical and scientific interests were wider in such countries. However, most differences showed little to no variation. Recently, Tao et al. (2022) provided a more comprehensive analysis, highlighting that across all dimensions of vocational interests studied, increased gender equality was associated with wider sex differences.

In summary, the literature review reveals that the link between sex differences and gender equality has been extensively studied in terms of overall scores and personality traits. Overall, there was no support for theories predicting a decline in sex differences with increased gender equality in these domains. However, there is support for the gender equality paradox in areas such as attitudes toward mathematics, reading, and personality traits. It is important to note that reductions in sex differences may still be observed in other domains not covered here. Despite the extensive literature on the topic, significant gaps remain. The first part on overall scores highlighted that intraindividual strengths remain substantially understudied and might relate to sex disparities in STEM. Meanwhile, the second part, focused on personality, indicated that a cross-temporal analysis could uncover a reduction or even an inversion of the paradox. In other words, both a focus on intraindividual strengths and cross-temporal analyses have been key areas of research for advancing our understanding of women's underrepresentation in STEM.

5.2 Intraindividual Academic Strengths and the Gender Equality Paradox

Given the gap identified in the literature review, the second part of this project aimed to explore sex differences in intraindividual strengths and the gender equality paradox in greater depth.

Study II

Sex Differences in Overall Scores

In Study I, I first analyzed sex differences in overall mathematics, reading, and science scores across five PISA waves. Boys consistently showed a slight advantage over girls in mathematics in each wave (Fig. 1). In reading, girls scored significantly higher than boys, although this gap narrowed slightly in the last two waves. For science, boys had a smaller advantage compared with mathematics, and in PISA 2012 and 2018, the difference was not statistically significant, indicating that sex differences in science scores nearly disappeared in these waves.

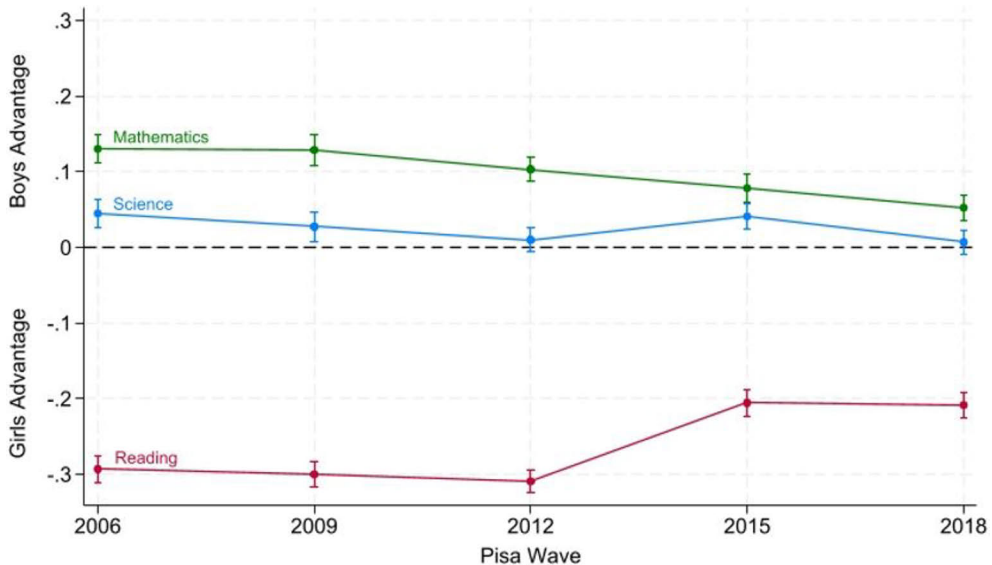


Figure 1. Wave-by-wave standardized mean sex differences (\pm Standard Error) in mathematics (green), reading comprehension (red), and science literacy (blue) PISA scores. Positive values indicate an advantage for boys, while negative values indicate an advantage for girls. (Figure reproduced from: doi-org.ezproxy.utu.fi/10.1177/09567976241271330).

Although wave-level sex differences are generally consistent, there is significant variation across countries within each PISA assessment. For example, examining sex differences in overall mathematics scores in PISA 2018, girls outperformed boys in 13 countries (16.9%), boys outperformed girls in 31 countries (40.3%), and 33 countries (42.9%) showed no statistically significant differences. Notably, there were no countries where sex differences in reading approached zero or favored boys in any of the PISA waves analyzed, highlighting a significant advantage for girls in this area. In science, patterns were similar to mathematics, showing substantial variation in sex differences across countries. In PISA 2015, 24 countries (34.8%) showed an advantage for boys, 20 (29%) an advantage for girls, and 25 (36.2%) showed no significant differences. Similar patterns were found across waves.

Sex Differences in Intraindividual Strengths

Sex differences in intraindividual strengths, unlike those in overall scores, demonstrated a remarkably stable pattern with minimal variation across countries. Overall, boys consistently scored higher in mathematics and science as intraindividual strengths, while girls consistently scored higher in reading as an intraindividual strength. This stability is evident, as across the five PISA waves analyzed, there was no country where girls outperformed boys in mathematics or where boys outperformed girls in reading when considering their respective intraindividual strengths in these subjects. Additionally, there were only a few instances (the U.S. in PISA 2006, Jordan in PISA 2009, and Lebanon in PISA 2015) where girls outperformed boys in science. In other words, sex differences in mathematics and reading are universal (i.e., found across countries), while differences in science are nearly universal.

Sex differences in intraindividual strengths are not only more stable than those in overall scores but also tend to be much larger (Fig. 2). For instance, in Belgium, in PISA 2015, sex differences in overall reading scores were around .3 SDs—a small-to-medium effect. However, in reading as an intraindividual strength, the difference exceeded 1 SD—a very large effect size by statistical standards. This pattern of larger sex differences in intraindividual strengths was consistent across countries and PISA waves.

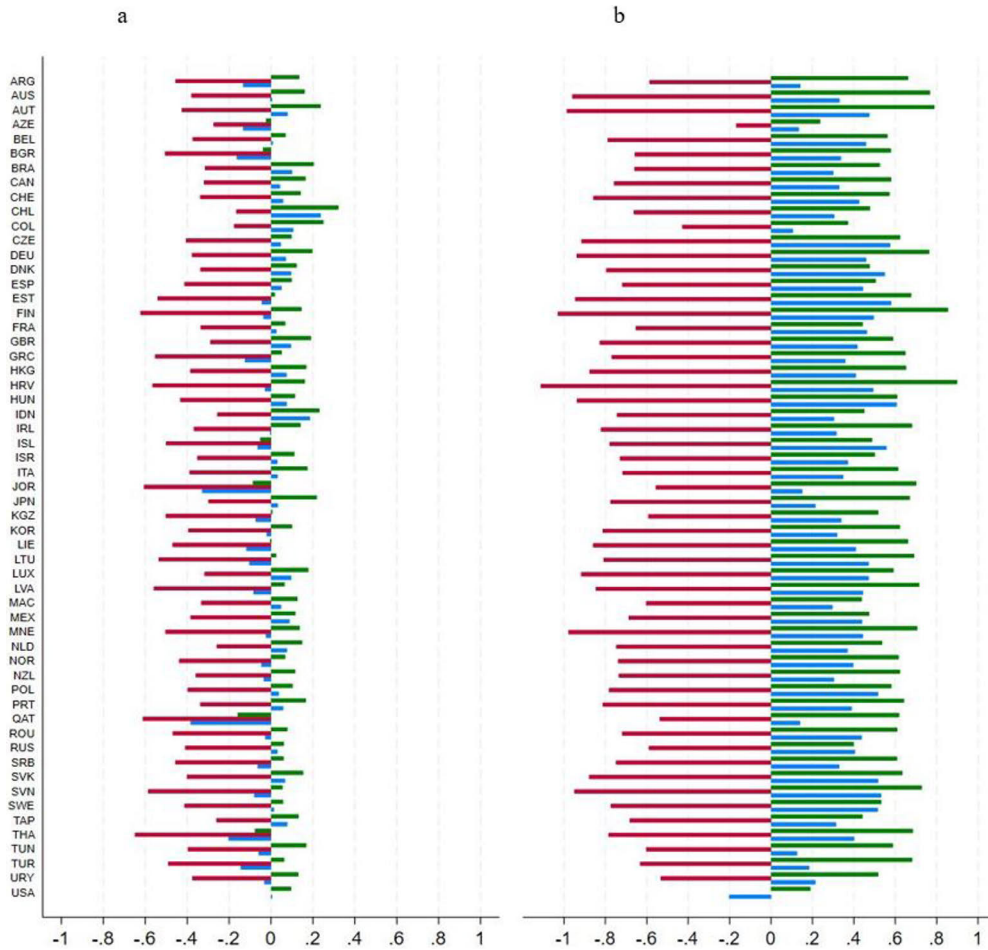


Figure 2. PISA 2006 standardized sex differences in mathematics (green), reading (red), and science (blue) as mean overall scores (a) and as intraindividual strengths (b). Values below 0 indicate an advantage for girls. (Figure reproduced from: doi.org.ezproxy.utu.fi/10.1177/09567976241271330).

The Gender Equality Paradox

Replicating previous research (Stoet & Geary 2018), no correlation between sex differences in mathematics as an intraindividual strength and GGI was found. Said differently, sex differences in this domain arise regardless of national-level gender equality.

However, sex differences in reading and science were consistently positively correlated with GGI (Fig. 3), implying that as gender equality increases, so do sex differences in reading and science as intraindividual strengths. With some fluctuations, results were replicated in the five PISA waves considered.

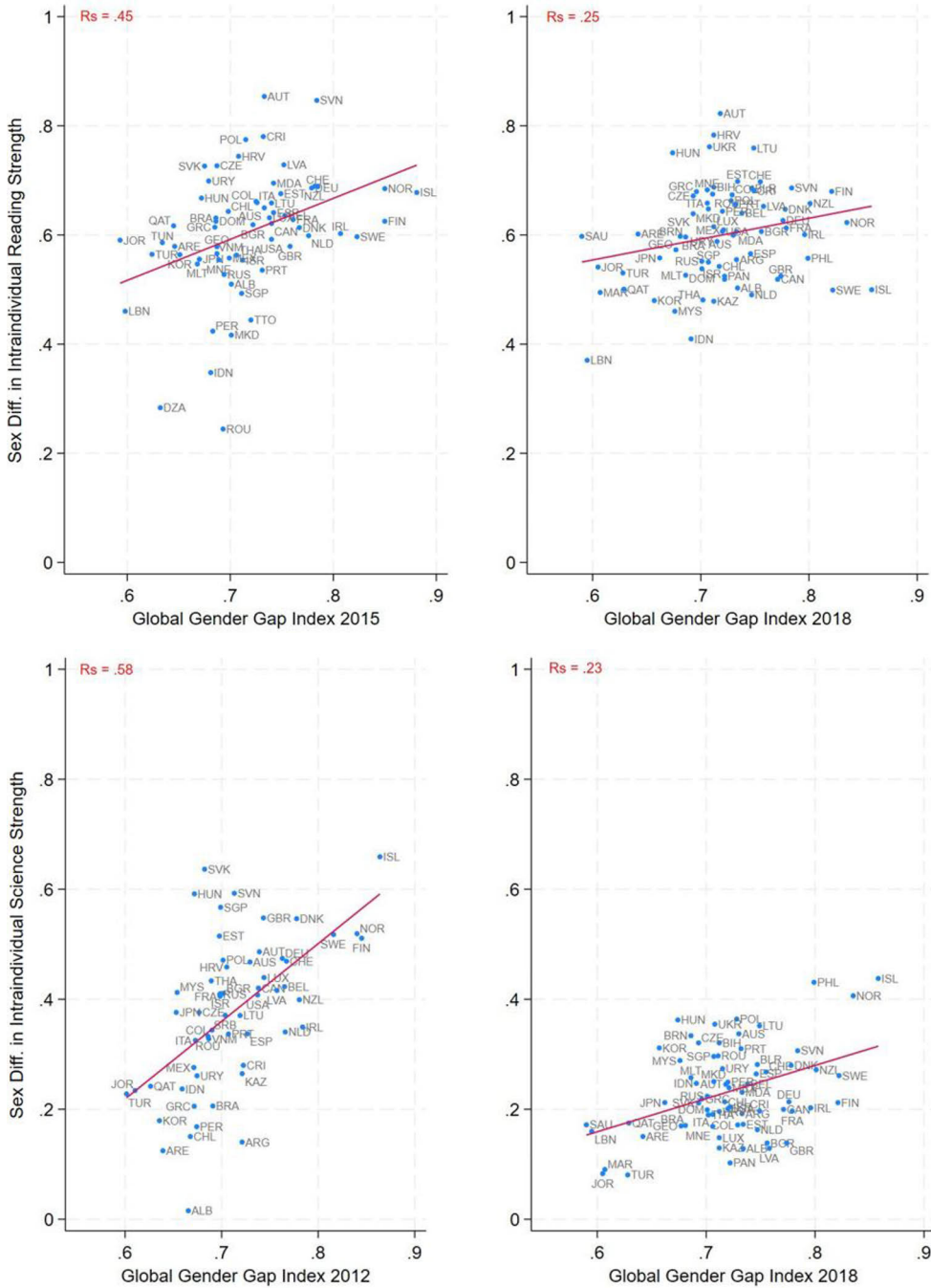


Figure 3. Relation between sex differences in reading (top) and science (bottom) as an intraindividual strength and the Global Gender Gap Index (GGGI). Highest and lowest coefficients across waves are reported. (Figure reproduced from: doi.org.ezproxy.utu.fi/10.1177/09567976241271330).

Similar results were found when using the wealth index. No statistically significant correlation emerged between country wealth and sex differences in mathematics as an intraindividual strength. However, higher country wealth was associated with larger sex differences in reading and science as intraindividual strengths. Notably, the relationship between country wealth and reading differences appeared stronger than that observed for the GGGI, whereas, for science, the association was more pronounced with the GGGI than with wealth.

The OLS regression models align with the results from the Spearman correlations. Generally, a 1 SD increase in GGGI scores in the null model (i.e., not controlling for wealth) was associated with a 0.35–0.38 SD increase in girls' advantage in reading as an intraindividual strength. For science, the coefficient was approximately 0.50 SDs (average across waves).

Wealth displayed similar results for reading, but not for science. Sex differences in reading increased with country wealth across all five PISA waves (coefficient ranged from 0.25 to 0.62 SDs). However, wealth was not significantly related to sex differences in science in PISA 2015 and 2018. Wealth explained, on average, more variation in sex differences in reading than in science, whereas GGGI showed the opposite pattern. Moreover, in the full model including both wealth and GGGI, the effect of GGGI disappeared or was substantially reduced when controlling for wealth. These results suggest, once again, that wealth shares a stronger association with sex differences in reading as an intraindividual strength, while GGGI relates more to differences in science.

In the final analysis, the sample was restricted to countries that participated in all five PISA waves ($N = 44$). First, I replicated the above analyses separately for each wave. Then, using longitudinal OLS with random (with GGGI-by-wave interaction) and country-fixed effects, I explored the gender equality paradox overtime.

OLS models run separately for each PISA wave largely replicated the findings from the full sample, except for 2018, where no significant relation was found between sex differences in reading as an intraindividual strength and both GGGI and WEALTH. While the effect size for the gender equality paradox appeared to decrease in the 2015 and 2018 PISA waves, longitudinal random-effects models—including interactions between GGGI (or wealth) and wave (time-variable)—did not yield statistically significant results. Similarly, longitudinal fixed-effects models were not statistically significant, showing no temporal relationship between GGGI or wealth and sex differences in reading or science as intraindividual strengths within countries, likely due to the high consistency of both GGGI and WEALTH over time.

Study III

Study III expands on the findings from Study II by examining sex differences in intraindividual strengths and the gender equality paradox across different PISA achievement levels.

Sex Differences across PISA Achievement Levels

Intraindividual Strengths

Sex differences in mathematics as an intraindividual strength favored boys across PISA achievement levels (Fig. 4). While some contrasts were statistically significant, ANOVA tests indicated that these differences remain largely consistent across achievement levels. Across countries, there was notable stability in the direction of sex differences, with only low achievers in China and the Philippines in PISA 2015 and 2018 showing no sex differences.

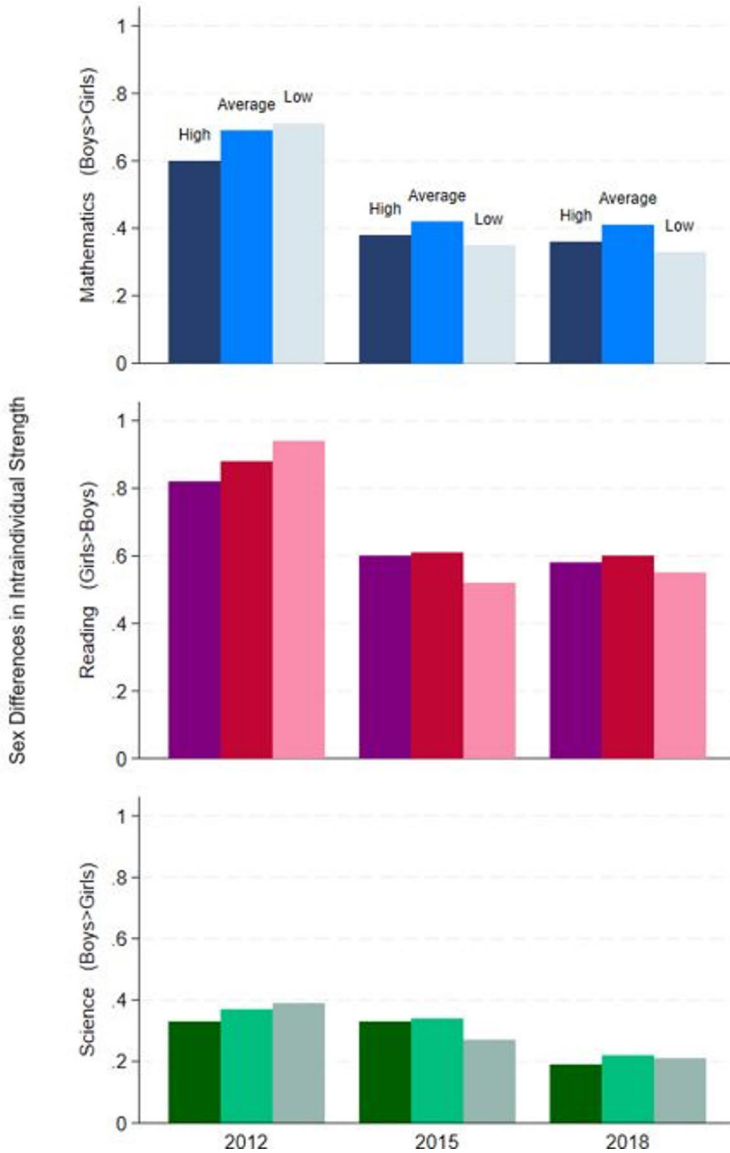


Figure 4. Wave-by-wave standardized mean sex differences in mathematics (blue), reading (red), and science (green) as intraindividual strengths among high, intermediate, and low PISA achievers.

For reading, the pattern was similar but instead favoring girls. Sex differences in this dimension are also significantly larger than those in mathematics. ANOVA tests showed that sex differences in reading as an intraindividual strength were consistent regardless of achievement levels. Testifying the substantial advantage that girls hold,

there were no countries where boys outperformed girls in reading or where sex parity was observed at any PISA achievement level and in any wave.

In science, sex differences again favored boys, although the magnitude was somewhat smaller than that observed for mathematics and remained stable across achievement levels. These differences were nearly universal, that is, present in almost all countries, waves, and PISA achievement levels, with only a few exceptions. For instance, a small advantage for girls was found in all PISA assessments for high achievers in Jordan (-0.11 to -0.15 SDs), and in PISA 2018 for low achievers in Malta (-0.15 SD) and Turkey (-0.10 SD).

Relative Ratio

Analyzing sex differences using relative ratios highlights the number of girls, compared with boys, with mathematics, reading, or science as the main intraindividual strength, rather than merely focusing on the magnitude of the differences.

Mathematics

Overall, more boys than girls had mathematics as their primary intraindividual strength across countries, waves, and achievement levels. Among high and average achievers, the ratio was approximately 3 to 1, decreasing to 2 to 1 among lower performers (average across PISA waves). The most significant disparity favoring boys was observed among low achievers in Jordan, where the ratio was 25 boys for every girl in PISA 2012. Conversely, there were countries where the ratio favored girls. For instance, in Thailand, a ratio of 1.89 girls for every boy was observed among high achievers in the same year.

Reading

In reading, the ratio consistently favored girls at every achievement level, with a striking 10.31 girls for every boy among high achievers in PISA 2012. Although this ratio decreased for other achievement levels and waves, it remained substantial. This outcome was also highest among high achievers and steadily declined as achievement levels decreased. Very few countries exhibited a ratio favoring boys; the most notable case was Qatar, where there were 3 boys for every girl with reading as their main intraindividual strength among low achievers.

Science

There were generally more boys than girls with science as their main intraindividual strength across achievement levels. However, boys' advantage in science was smaller than that observed in mathematics. Among high achievers, there were, on average, 1.37 boys for every girl. The ratios for average and low achievers were 1.53 and 2.13

boys per girl, respectively. The largest disparity occurred among lower achievers in Qatar (14:1), while Jordan had 4.41 girls for every boy among high achievers, but few countries showed a ratio favoring girls.

Percentages

The relative ratio pattern can also be traced by examining the percentage of boys and girls who have mathematics, reading, or science as their main intraindividual strength (Fig 5). The findings are consistent with the previous analyses, reflecting similar trends across PISA achievement levels.

Among *high achievers*, 42% of girls had reading as their main intraindividual strength, followed by science and mathematics at 30% and 28%, respectively (average across PISA waves). For boys, mathematics was the main strength for 46%, followed by science at 41% and reading at only 13%.

The results were even more striking among *average achievers*, where 55% of girls had reading as their main intraindividual strength, while rates of 23% and 22% were noted for science and mathematics, respectively. Meanwhile, boys showed the main intraindividual strengths in mathematics (43%), followed by science (38%), and, lastly, reading (19%).

Among *low achievers*, reading was the main strength for 44% of girls, mathematics for 28%, and science for 28%. Boys, however, revealed strengths in mathematics (46%), with lower figures for science (38%) and reading (16%).

In summary, boys were 52%, 80%, and 48% (relative percentage difference) more likely to have mathematics or science as their main strengths across high, average, and low achievement levels, respectively. Conversely, girls were 231%, 189%, and 175% more likely to have reading as their main strength at these same levels.

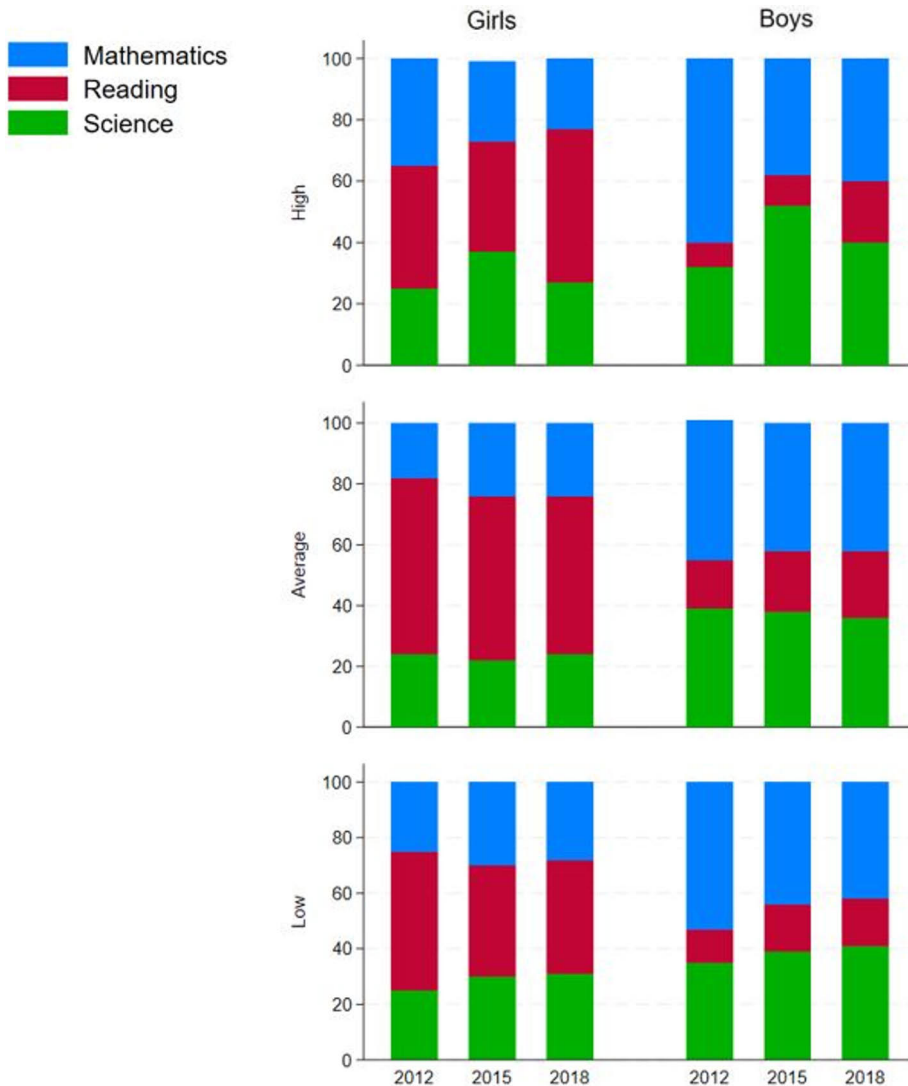


Figure 5. Wave-by-wave percentages of girls (left) and boys (right) with either mathematics (blue) reading (red) or science (green) as main intraindividual strength among high (top) intermediate (middle) and low (bottom) PISA performers. See S7 Table for the country-level results.

Correlations among Sex Differences in Intraindividual Strengths

An aspect that remains unclear is whether sex differences in intraindividual strengths exhibit the same patterns as differences in overall scores in mathematics, reading, and science. To investigate this trend, intraindividual strengths were correlated. As stated above, at the country level, sex differences in overall mathematics and reading

scores tend to show an inverse relationship, indicating that higher sex differences in mathematics were associated with lower sex differences in reading.

Interestingly, unlike the patterns seen in overall scores, a positive correlation was observed between sex differences in mathematics and reading as intraindividual strengths at the country level across all PISA achievement levels (Fig. 6). This outcome suggests that as the sex difference in mathematics as an intraindividual strength increases, so does the difference in reading. Although there were slight fluctuations in the strength of this relationship, it remained generally stable across all PISA achievement levels and waves.

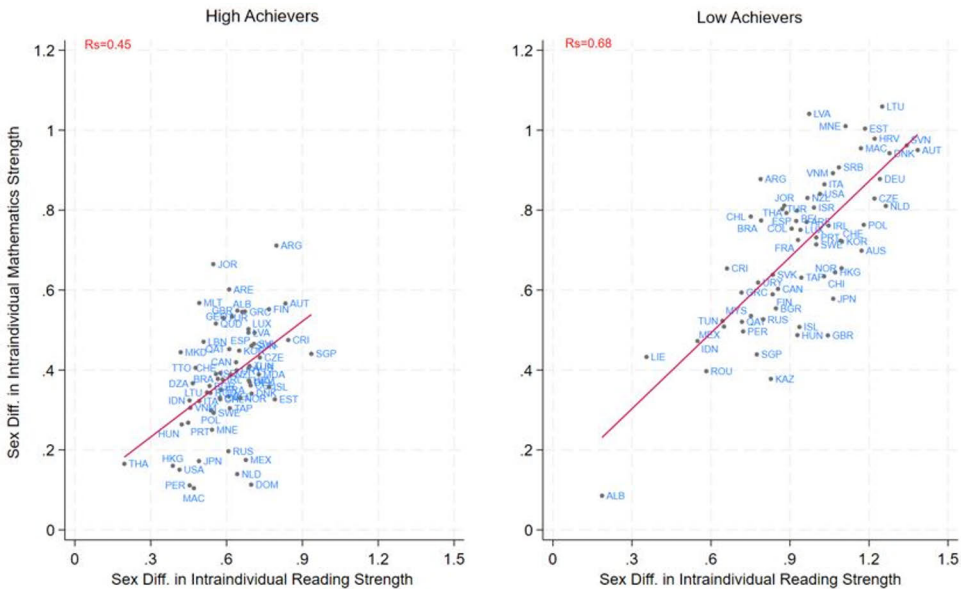


Figure 6. Correlation between sex differences in mathematics (boys > girls) and reading (girls > boys) as intraindividual strengths among high (left panel) PISA 2015 and low (right panel) PISA 2012 achievers. Highest and lowest coefficients across waves and achievement levels are reported.

A similar pattern was observed in the correlation between sex differences in reading and science as intraindividual strengths. That is, as girls' advantage in reading as an intraindividual strength increased at the country level, boys' advantage in science also increased. However, the strength of this relationship reduced from earlier to more recent PISA waves at each achievement level. Despite this decline, the correlation remained statistically significant throughout.

Surprisingly, I found a negative correlation between sex differences in mathematics and science as intraindividual strengths (Fig. 7). This result suggests that as the former increases, the latter tends to decrease at the country level. With the

exception of low and average achievers in PISA 2012, this correlation remained consistent across all PISA waves and achievement levels.

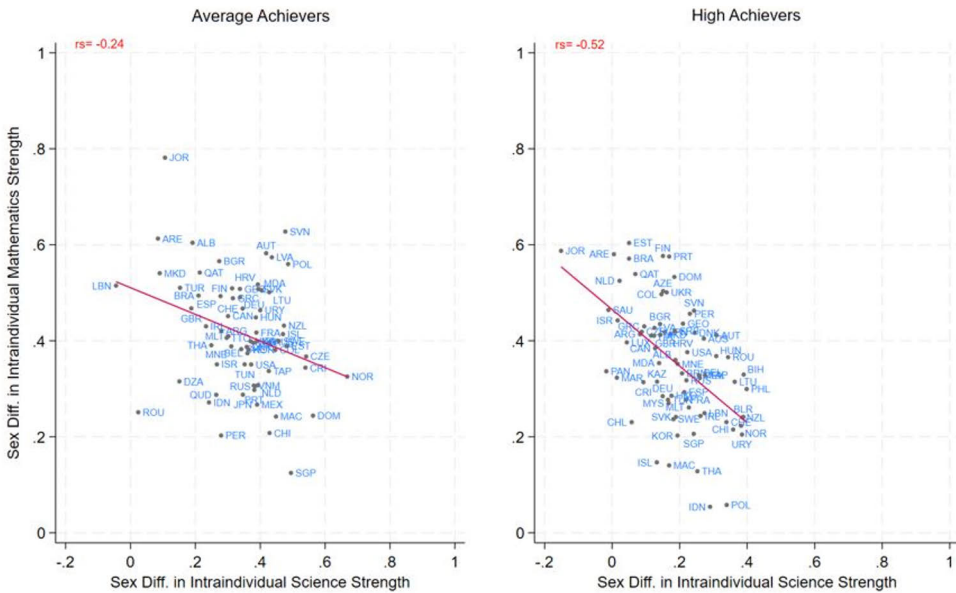


Figure 7. Correlation between sex differences in mathematics and science as intraindividual strengths (boys > girls), among intermediate (left panel) and high (right panel) achievers for PISA 2015 and 2018, respectively. Highest and lowest coefficients across waves and achievement levels are reported.

The Gender Equality Paradox

In the final part of Study III, I examined whether the gender equality paradox persisted across different levels of PISA achievement. The results closely mirrored those from Study II. Overall, there was no significant correlation between sex differences in mathematics as an intraindividual strength and GGI at any achievement level.

In contrast, the gender equality paradox for sex differences in reading and science as intraindividual strengths was largely consistent across achievement levels, with some exceptions (Fig. 8). For low achievers in PISA 2015 and average and high achievers in PISA 2018, the correlation between reading differences and GGI was not significant, likely due to Iceland and Sweden acting as outliers. Removing these countries from the analysis resulted in correlations aligned with the gender equality paradox.

Similarly, a positive relationship between sex differences in science and GGI scores was found, except for high and average achievers in PISA 2018, where Iceland again influenced the null results. Excluding Iceland from the analysis yielded correlations consistent with other findings. Notably, coefficients were generally

higher for science than for reading, suggesting once again that GGGI is more closely related to sex differences in science than in reading as intraindividual strengths.

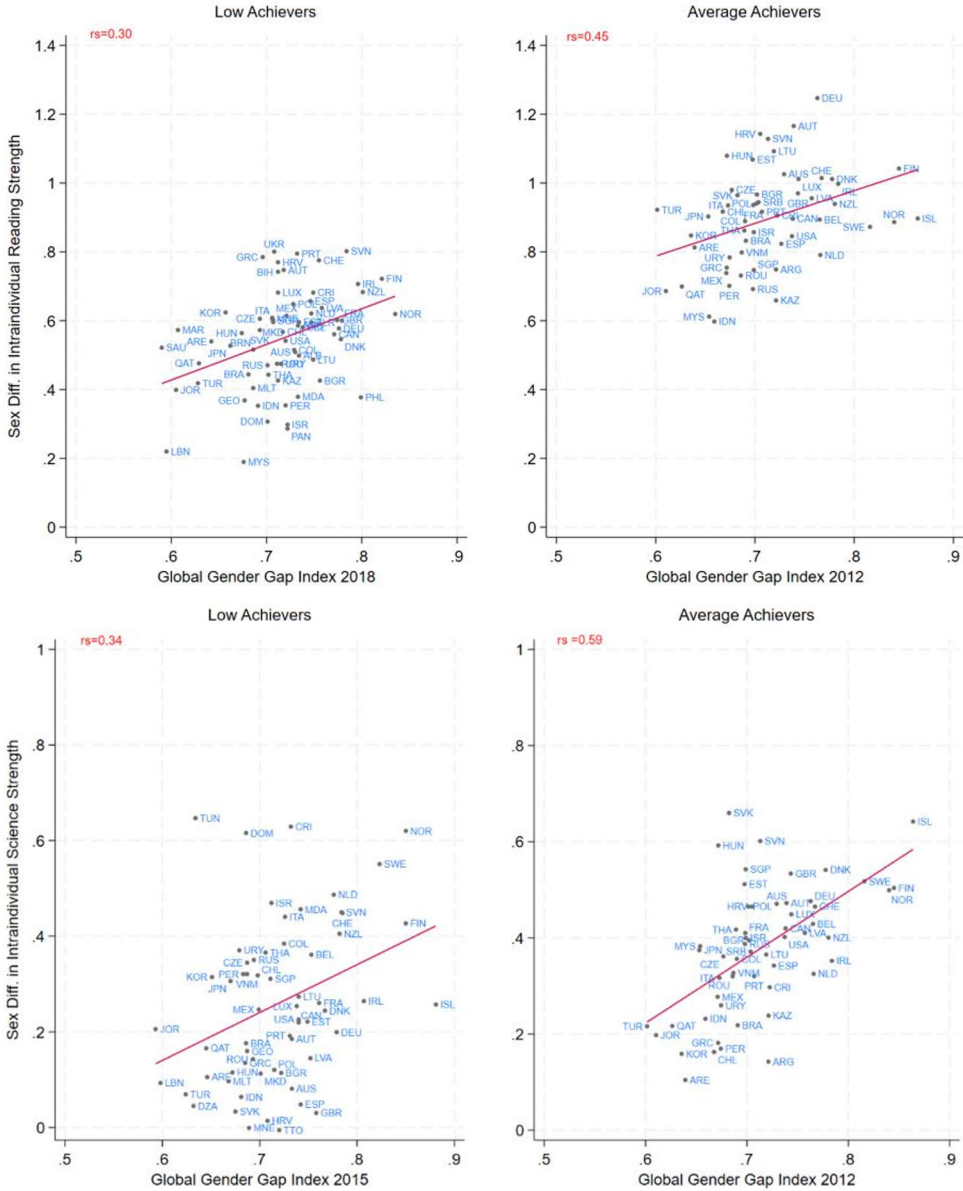


Figure 8. Sex differences in science (boys > girls) and reading (girls > boys) as an intraindividual strength and their relation with GGGI across PISA achievement levels. Highest and lowest significant rs across waves and achievement levels are reported.

Lastly, OLS models support the correlation, indicating that sex differences in reading and science generally increase with gains in national gender equality. The coefficients were more stable than the correlation and were stronger for science than for reading. Specifically, a 1 SD increase in sex differences in reading was associated with a 0.30 SD increase in reading as an intraindividual strength at each PISA achievement level. The coefficient for science was approximately 0.40 SD. Iceland and Sweden acted as outliers only for sex differences in reading among high and average achievers in PISA 2018.

Lastly, examining sex differences alone may obscure which group—boys or girls—is driving the paradox. One group may experience a sharper increase or decrease in response to greater gender equality, leading to widening sex differences. To investigate this issue, I analyzed whether boys, girls, or both were contributing to the paradox. Specifically, I included an interaction term between the mean national performance of boys and girls in each intraindividual strength and the GGI score. Interestingly, the interaction models revealed that, at the country level, the group with an overall advantage tended to improve its intraindividual performance as the GGI increased, while the group with an overall disadvantage tended to exhibit poorer intraindividual performance. For instance, among low achievers in PISA 2012 (Fig. 9), girls' intraindividual strength in reading increased, while boys' intraindividual strength decreased with gains in gender equality (sex by GGI, $\beta = -.14$, $p = .01$).

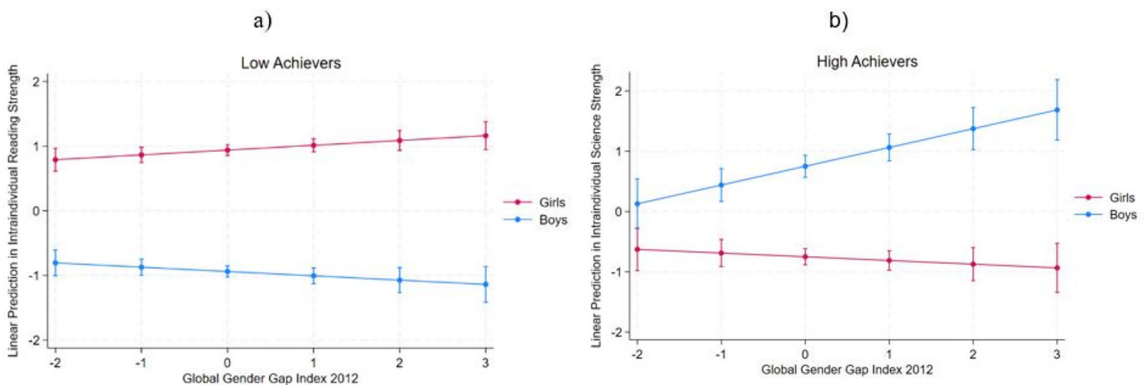


Figure 9. Linear prediction for boys and girls low achievers in reading (panel a) and high achievers in science (panel b) as intraindividual strengths for in PISA 2012.

6 Discussion

This study examined a specific aspect of the gender equality paradox in STEM education: intraindividual academic strengths. It provides both theoretical and methodological contributions by integrating current literature on the paradox and analyzing sex differences in intraindividual strengths across gender equality levels, time periods, and academic achievement percentiles.

6.1 Summary of the Findings

A Review of Sex Differences and Their Relation to Gender Equality

A systematic narrative review of the literature on sex differences in mathematics, reading, and science, as well as personality traits, and their associations with gender equality indicators, reveals that differences generally either increase or remain stable in more gender-equal countries. Specifically, sex differences in mathematics do not correlate with gender equality indices, while differences in reading often show a positive correlation. As the sex gap in reading increases, women may have even more opportunities to excel in fields outside STEM (Wang et al., 2013). Studies on sex differences in science and aggregate scores are limited and firm conclusions cannot be drawn. Notably, sex differences in mathematics attitudes and anxiety align with the gender equality paradox—they increase with greater gender equality.

Stoet and Geary (2018) analyzed sex differences in intraindividual strengths in mathematics, reading comprehension, and science literacy. They observed that girls typically have reading as an intraindividual strength, while it is mathematics or science among boys. These differences were also wider in more gender-equal countries. However, their study was the only one to focus on intraindividual strengths, which could be a significant determinant of STEM participation (Dekhtyar et al., 2018). Despite its potential importance, this area of research has been largely neglected, and some scholars have also questioned the replicability of Stoet and Geary's findings (Richardson et al., 2020).

However, the gender equality paradox has been extensively studied in the context of personality, with a consistent trend observed across various models, including the Big Five, HEXACO, basic human values, and vocational interests frameworks. Sex differences in traits such as neuroticism and agreeableness (Big Five) and emotionality (HEXACO) increase with improvements in national-level gender equality. Similarly, differences associated with men favoring power, achievement, and stimulation and women favoring benevolence and universalism (basic human values model), as well as sex differences in vocational interests, follow a similar pattern. Notably, Fors Connolly and colleagues (2020) observed that sex differences in basic human values decrease over time, highlighting the importance of examining the gender equality paradox from a temporal perspective.

In summary, the review highlighted two significant gaps in the literature. First, a potentially critical aspect of the gender equality paradox in STEM—namely, intraindividual academic strengths in mathematics, reading, and science—has been substantially understudied. Second, research on sex differences in personality traits suggests that the patterns among men and women may be converging over time, emphasizing the importance of incorporating temporal analyses when examining the paradox.

Sex Differences in Intraindividual Strengths across Gender Equality Scores, Time, and Achievement Levels

The second part of this dissertation built upon the findings from the literature review and examined sex differences in intraindividual strengths in mathematics, reading, and science using PISA data. Study II revealed that boys generally scored higher in mathematics and science as intraindividual strengths, while girls scored higher in reading. These sex differences were not only more stable but also generally much larger than those in overall scores. Additionally, differences in reading and science as intraindividual strengths increased with gains in gender equality, as measured by the GGGI. This association was consistently found across all five PISA waves analyzed, without significant changes over time. However, sex differences in mathematics were independent of country-level gender equality.

Exploring sex differences in intraindividual strengths at various PISA achievement levels, Study III expands the findings of Study II. The same pattern emerged: mathematics and science were more likely to be strengths for boys, while reading was more likely to be a strength for girls. These trends remained stable across achievement levels, although some fluctuations were noted. Similar results were observed when examining both the ratio of girls to boys and the percentage of girls and boys with mathematics, reading, or science as their main strength. Additionally, sex differences in reading were positively correlated with those in mathematics and

science, though mathematics and science differences showed a surprising negative relationship. This unexpected correlation could stem from adolescents who are interested in mathematics/science making extra curriculum choices that favor one or the other, but this hypothesis remains to be tested. Overall, the results suggest that not only are sex differences in intraindividual strengths larger and more stable than differences in overall scores as evidenced by Study II, they also follow distinct patterns at the country level. Lastly, even in Study III, the gender equality paradox in reading and science as intraindividual strengths persisted across PISA achievement levels, while sex differences in mathematics as an intraindividual strength were never significantly correlated with the GGGI.

6.2 Implications

The results of this dissertation have significant implications for addressing the persistent underrepresentation of women in STEM, even in wealthy, gender-equal countries that have made substantial investments in minimizing this disparity. The implications pertain not only to intraindividual strengths but also to the gender equality paradox within this domain.

Intraindividual Strengths

According to expectancy-value theory (Eccles, 1983), intraindividual strengths influence educational choices. Given the observed sex differences in these strengths, it could be expected that more boys would choose STEM fields, while more girls would prefer the humanities and social sciences. This trend would be especially pronounced in wealthier and more gender-equal countries, where individuals can separate their educational choices from the expected long-term economic return of their selected fields (Stoet & Geary, 2018). This is not to suggest that intraindividual strengths are the sole cause of women's underrepresentation in STEM; other factors, such as stereotypes (Napp & Breda, 2022), could also play a role.

The debate about the underrepresentation of women in STEM fields mostly focuses on educational subjects tied to white-collar occupations, which are generally cognitively demanding and associated with a high income and status. However, PISA scores are strongly correlated with the intelligence quotient (IQ), with higher PISA scores typically reflecting higher levels of cognitive ability (Lynn & Mikk, 2009; Rindermann, 2007). Thus, the fact that sex differences in intraindividual strengths are found across PISA achievement levels suggests that these strengths influence more than just educational pathways linked to high-skilled, white-collar professions. The implication is that intraindividual strengths also relate to male-dominated, less cognitively demanding, blue-collar fields (e.g., mechanical work,

carpentry) (Prix & Kilpi-Jakonen, 2022), which adolescent boys show a significantly higher inclination to pursue compared with girls (Stoet & Geary, 2022). In other words, based on how consistent sex differences in intraindividual academic strengths are across PISA achievement levels, we should not only expect more boys than girls to follow white-collar STEM fields but also a higher number of boys than girls to enter blue-collar occupations. However, sex disparities in blue-collar occupations have received less attention from scholars. Further research should focus on the determinants of these careers, even though they are less closely associated with status, and intraindividual strengths offer a valid starting point.

Lastly, the much larger magnitude of sex differences in intraindividual strengths compared to differences in mean scores highlights the importance of moving beyond these mean scores when exploring cognitive factors linked to the underrepresentation of women in STEM.

The Gender Equality Paradox

Social constructivist theories propose that sex differences would diminish as more gender equality is achieved. By contrast, other theories (e.g., evolutionary frameworks) have predicted the opposite trend. The results of this dissertation, along with a recent meta-analysis (Herlitz et al., 2024), indicate that sex differences typically remain stable or are exaggerated as country-level gender equality increases. Moreover, the gender equality paradox, at least in relation to intraindividual strengths, appears to be consistently observed across time, addressing previous concerns about its replicability (Richardson et al., 2020).

Taken together, these findings challenge social constructivist theories and seem to be more in line with alternative explanations. Indeed, many sex differences have declined over time and even been reversed in some cases as a result of increasing economic opportunities for women. For instance, in the early 20th century, men outnumbered women in fields such as medicine and veterinary medicine (Winston, 1935). However, today, in wealthier and more gender-equal countries, medical school enrollment has reached parity (Pelley & Carnes, 2020), while veterinary medicine is now female-dominated (Lofstedt, 2003). In other words, the theoretical implication of the paradox is that, when examining sex differences and their changes with increased gender equality, a more comprehensive framework is needed—one that accounts for both the social environment and the evolved, intrinsic characteristics of men and women. While some sex differences can still be expected to reduce or even disappear, this expectation must be balanced against the potential evolutionary origins of the specific traits studied. Note, however, that social norms are still important for the full expression of sex differences (see below).

The theoretical implications outlined above also extend to policymaking. Many current policies are influenced by social constructivist views, which suggest that sex disparities in STEM fields are primarily the result of socialization processes and the internalization of sex roles (Hervías Parejo & Radulović, 2023). Consequently, these policies often interpret the underrepresentation of women as a sign of enduring societal forces that discourage them from pursuing these careers. However, the gender equality paradox in intraindividual strengths raises important considerations; it proposes that the absolute convergence between men and women in STEM might also be shaped by individual characteristics that, on average, vary between people of both sexes—such as intraindividual strengths. In other words, efforts to achieve parity between men and women in STEM should consider that career choices are shaped by a complex interplay of factors, some of which may not be easily addressed through policy interventions but are more closely related to individual choices. In this vein, rather than proposing equity targets, policies could focus on expanding opportunities and supporting individuals in pursuing careers aligned with their strengths and interests while ensuring these are not shaped by socialization pressures.

Regardless, the novelty of the intraindividual strengths concept—which is yet to be considered in policymaking—suggests new opportunities to encourage more women to pursue STEM careers. Targeted interventions could be designed for girls who show intraindividual strengths in mathematics or science, providing them with mentorship, hands-on experiences, and support to build their interest and confidence in these fields. Similarly, boys who exhibit strengths in reading could benefit from tailored programs to nurture their potential for more verbally oriented domains.

6.3 Possible Explanations of the Gender Equality Paradox

Some researchers argue that the gender equality paradox could stem from an increase in stereotypes in more gender-equal countries. For instance, Napp and Breda (2020) argue that the stereotype linking mathematics to boys and reading to girls is more pronounced in these countries. Stereotypes may indeed still influence women's underrepresentation in STEM. For example, Dekhtyar et al. (2018) found fewer women in STEM than what is expected based on intraindividual strengths. However, the stereotype explanation is likely not substantive. Attitudes toward women in education tend to be more supportive in gender-equal countries, where acceptance of gender equality is also higher (Donnelly et al., 2016; Stoet & Geary, 2020). Additionally, a substantial body of literature suggests that sex stereotypes are often accurate, reflecting genuine sex differences rather than distorted perceptions—a phenomenon known as stereotype accuracy (A. Eagly & Hall, 2024; Jussim et al., 2015).

Another potential explanation for the gender equality paradox in intraindividual strengths could be data quality issues stemming from cultural differences in how survey questions are interpreted, potentially leading to a "Simpson effect" (Berggren & Bergh, 2023). In other words, there might be variations in how questions are understood across countries and this might bias the results. While such an effect could occur for other measures, it is unlikely to be the case for PISA. Specifically, PISA employs a robust structure for data collection to ensure reliability and consistency, significantly minimizing the risk of interpretative discrepancies. Test items are extensively piloted in diverse linguistic contexts and they are scrutinized to identify and eliminate potential biases (OECD, 2009). This rigorous process ensures that the constructs being measured are comparable across all participating countries. Additionally, the data in Studies II and III were standardized at the country level, aligning variable scales across nations and further reducing the likelihood of a Simpson effect.

Ilmarinen and Lönqvist (2024) recently argued that the gender equality paradox in intraindividual strengths might be a statistical artifact. They claim that using algebraic differences between boys' and girls' means in bivariate analyses is problematic, as the results could be driven by correlations with gender equality indicators affecting either or both groups' means. Based on their analysis, they suggest that the gender equality paradox may arise from both men and women diverging equally (with one increasing and the other decreasing at similar rates) or from one sex diverging more than the other. They argue that a true gender equality paradox occurs only when both sexes diverge equally. However, with research on the paradox, a question that emerges is whether sex differences increase in more gender-equal countries, regardless of whether men, women, or both contribute. Therefore, even if only one group improves relative to the other in response to gender equality, sex differences will still increase, constituting a paradox. This dissertation examined how sex differences in intraindividual strengths change with increased gender equality, rather than isolating which group drives the phenomenon. Nevertheless, interaction effects from Study III suggest that the group with higher intraindividual scores in a particular area—such as girls in reading—tends to benefit more from increased gender equality, while the opposite effect occurs for the less-advantaged group. In other words, both boys and girls diverge more as gender equality increases.

Drawing on expectancy-value and evolutionary theories, this dissertation suggests that, especially in more gender-equal countries, intraindividual academic strengths contribute to the underrepresentation of women in STEM. In these countries, economic prosperity and social attitudes such as individualism and self-expression allow individuals to pursue personal preferences and reinforce abilities, including intraindividual strengths. Accordingly, gender equality and economic

development correlate highly in gender-equal countries but not in less gender-equal ones. Plots of the GGGI and wealth data from Study II show that wealthy but less gender-equal countries display smaller sex differences in intraindividual strengths than wealthy but more gender-equal countries. This result suggests that social attitudes significantly relate to sex differences in intraindividual strengths but only once a substantial level of national wealth is accumulated. In other words, social attitudes are still relevant for the full expression of sex differences along with economic prosperity. These attitudes might relate to social views toward girls' and women's education or, more broadly, any attitude that may enhance individuals' freedom to make choices (Stewart-Williams & Halsey, 2021). Note that, although academic skills like reading and mathematics are evolutionarily novel—meaning they did not specifically evolve—they still draw on cognitive systems shaped by sexual selection. This may explain why intraindividual strengths follow patterns consistent with the gender equality paradox. For instance, mathematics relies on spatial abilities where men often show an advantage, possibly due to the demands of hunting tasks in hunter-gatherer societies. These abilities are rooted in the "folk physics" domain, an innate human understanding of basic physical phenomena (Halpern et al., 2007). Similarly, reading skills depend on cognitive processes linked to language acquisition, which may have been especially important for women due to their ancestral, primary role as caregivers (Geary, 2024).

6.4 Limitations

The systematic narrative approach used in the first part of this study did not allow for a thorough re-analysis of the existing literature, hence the conclusions may be affected by the authors' interpretations. Nevertheless, a recent meta-analysis (Herlitz et al., 2024) has addressed this limitation, reaching results that substantially align with those described in the literature review. This adds further support for the conclusions presented in Study I, reinforcing them with evidence from a more rigorous and comprehensive evaluation of existing studies.

Studies II and III are correlational, thus subject to typical limitations of correlational studies, such as the inability to draw causal inferences. However, the research question on the gender equality paradox mainly focuses on whether sex differences increase or decrease with gender equality, rather than suggesting that gender equality is the root cause of the paradox. That is, the observed patterns may result from a combination of biological, cultural, and economic factors that interact in complex ways and are captured by gender equality indicators. One way to address this limitation is to use register-based data from the Nordic, and other, countries to examine how sex differences in intraindividual strengths respond to specific environmental changes. These datasets span several decades and enable robust

longitudinal analyses of how such differences may be shaped by historical shocks and broader societal developments.

Additionally, PISA data did not allow a direct test of expectancy-value theory, as data on student enrolment intentions were unavailable for most waves. Nonetheless, recent research supports the idea that intraindividual strengths significantly predict career paths. For instance, Stoet and Geary (2022) found that boys tend to aspire more to "things-oriented" fields, while girls aspire more to "people-oriented" fields. This difference widens in gender-equal countries and resembles the pattern observed for intraindividual academic strengths.

Further, some critics argue that GGI oversimplifies gender equality, failing to capture its multidimensionality. While GGI has its limitations (e.g., disregarding women's advantages over men), it aligns closely with social constructivist theories (Stoet & Geary, 2019). Moreover, the findings presented here are highly consistent across both PISA waves and achievement levels, indicating that GGI captures systematic differences between countries effectively. Given the strong relationship between GGI and wealth in more gender-equal countries, these systematic differences likely relate to economic development and social norms that enhance individuals' freedom of choice.

Lastly, it is possible that the sex differences in intraindividual strengths reported in Studies II and III may have been either over or underestimated. PISA is a low-stakes assessment for students (i.e., it does not affect their grades), and there may be systematic differences between boys and girls in how they approach the test. These potential differences in effort during test-taking could exacerbate sex differences in intraindividual strengths. Meanwhile, the methodology employed by PISA might conceal the true magnitude of sex differences. For instance, Wittmann (2005) noted that when computing sex differences using raw PISA scores, these are nearly twice as large as those derived from imputed scores.

6.5 Future Research

This study assessed intraindividual strengths based on the assumption that they are relevant to educational choices, although this assumption was not directly tested. Sex differences in intraindividual strengths have also mainly been studied in high school students and it remains unclear if the same patterns are also found at later life stages. Further, previous research suggests that intraindividual strengths may predict career choices more effectively than overall scores in mathematics, reading, and science, but no study has assessed whether this hypothesis holds true. Additionally, intraindividual strengths are both theoretically and statistically related to "ability tilt," which is calculated by subtracting one domain-specific score (e.g., reading) from another (e.g., mathematics). Unlike ability tilt, intraindividual strengths

measure deviations from an individual's average performance, which aligns more closely with expectancy-value theory. However, research comparing the predictive power of ability tilt and intraindividual strengths in relation to career paths is still missing. Lastly, there is a lack of understanding regarding which processes, societal and/or biological, might influence intraindividual strengths and contribute to cross-country variation.

Considering all the abovementioned issues, future research should prioritize replicating the findings of this dissertation using alternative datasets, such as PIAAC. It would also be valuable to compare whether overall scores in mathematics, reading, and science; intraindividual strengths; or ability tilt better predict career choices. Other important research directions include examining how sex differences in these domains are linked to women's underrepresentation in STEM at the national level. Furthermore, examining variations between countries would also be valuable for understanding which cultural processes influence intraindividual strengths and related sex differences. This approach could offer valuable insights into the long-debated question of how specific academic abilities are linked to occupational sex differences, especially in STEM fields.

As another direction for future research, efforts could be made to reconcile findings on sex disparities in STEM participation in education and in the job market. While a substantial body of evidence highlights a persistent, or even increasing, gap between men and women pursuing STEM education (Blackburn, 2017; Stoet & Geary, 2018), some argue that these disparities are diminishing in the workforce (Almstedt Valldor & Halldén, 2023).

6.6 Conclusions

The findings of this dissertation suggest that intraindividual academic strengths might play a significant role in shaping career trajectories. Accordingly, shifting the focus from sex differences in overall scores in mathematics, reading, and science to differences in intraindividual strengths may offer new insights into the underrepresentation of women in STEM fields. Sex differences in intraindividual strengths are not only more stable than those observed in overall scores but are also considerably larger in magnitude.

Furthermore, the gender equality paradox—observed in intraindividual strengths as well as in many other domains (Herlitz et al., 2024; Vishkin, 2022)—challenges the social constructivist view that sex differences are solely the result of socialization processes. The fact that many of these differences either increase or remain stable as gender equality improves suggests the need for a more comprehensive perspective on their origins, one that considers both the evolutionary history of men and women, and the specific social contexts in which these differences are expressed.

The paradox also raises important questions for policymakers about how to address the underrepresentation of women in STEM. While efforts to encourage mathematically talented women to pursue these careers should certainly continue, they should also be weighed against the fact that certain individual characteristics—which, on average, differ between men and women—may limit the extent to which complete sex parity in STEM fields is achievable.

List of References

- Almstedt Valldor, A., & Halldén, K. (2023). Skills and occupational sex segregation in Europe. In M. Tählin (Ed.), *A Research Agenda for Skills and Inequality* (pp. 65–83). Edward Elgar Publishing. <https://doi.org/10.4337/9781800378469.00011>
- Anghel, B., Rodríguez-Planas, N., & Sanz-de-Galdeano, A. (2019). *Culture, Gender, and Math: A Revisitation*.
- Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational Assessment, 27*(3), 197–205.
- Baker, D. P., & Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education, 91*–103.
- Barone, C., & Assirelli, G. (2020). Gender segregation in higher education: An empirical test of seven explanations. *Higher Education, 79*(1), 55–78. <https://doi.org/10.1007/s10734-019-00396-2>
- Berggren, M., & Bergh, R. (2023). *Simpson's Gender-Equality Paradox*. <https://doi.org/10.31234/osf.io/mfhyw>
- Bergold, S., Wendt, H., Kasper, D., & Steinmayr, R. (2017). Academic competencies: Their interrelatedness and gender differences at their high end. *Journal of Educational Psychology, 109*(3), 439–449. <https://doi.org/10.1037/edu0000140>
- Bernstein, B. O., Lubinski, D., & Benbow, C. P. (2019). Psychological Constellations Assessed at Age 13 Predict Distinct Forms of Eminence 35 Years Later. *Psychological Science, 30*(3), 444–454. <https://doi.org/10.1177/0956797618822524>
- Blackburn, H. (2017). The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017. *Science & Technology Libraries, 36*(3), 235–273. <https://doi.org/10.1080/0194262X.2017.1371658>
- Bleidorn, W., Arslan, R. C., Denissen, J. J., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem—A cross-cultural window. *Journal of Personality and Social Psychology, 111*(3), 396.
- Bos, W., Wendt, H., Ünlü, A., Valtin, R., Euen, B., Kasper, D., & Tarelli, I. (2012). Leistungsprofile von Viertklässlerinnen und Viertklässlern in Deutschland [Proficiency profiles of fourth grade students in Germany]. In *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich [IGLU 2011. International Comparison of Reading Competencies of Primary School Children in Germany]*. Waxmann.
- Breda, T., & Napp, C. (2019). Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences, 116*(31), 15435–15440. <https://doi.org/10.1073/pnas.1905779116>
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*(4), 676.
- Cech, E. A. (2013). The Self-Expressive Edge of Occupational Sex Segregation. *American Journal of Sociology, 119*(3), 747–789. <https://doi.org/10.1086/673969>
- Charles, M., & Bradley, K. (2009). Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries. *American Journal of Sociology, 114*(4), 924–976. <https://doi.org/10.1086/595942>

- Charles, M., Harr, B., Cech, E., & Hendley, A. (2014). Who likes math where? Gender differences in eighth-graders' attitudes around the world. *International Studies in Sociology of Education*, 24(1), 85–112. <https://doi.org/10.1080/09620214.2014.895140>
- Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81(2), 322–331. <https://doi.org/10.1037/0022-3514.81.2.322>
- Dekhtyar, S., Weber, D., Helgertz, J., & Herlitz, A. (2018). Sex differences in academic strengths contribute to gender segregation in education and occupation: A longitudinal examination of 167,776 individuals. *Intelligence*, 67, 84–92.
- Devine, A., Fawcett, K., Szücs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions*, 8(1), 33. <https://doi.org/10.1186/1744-9081-8-33>
- Donnelly, K., Twenge, J. M., Clark, M. A., Shaikh, S. K., Beiler-May, A., & Carter, N. T. (2016). Attitudes Toward Women's Work and Family Roles in the United States, 1976–2013. *Psychology of Women Quarterly*, 40(1), 41–54. <https://doi.org/10.1177/0361684315590774>
- Eagly, A. H., & Mitchell, A. A. (2004). *Social role theory of sex differences and similarities: Implications for the sociopolitical attitudes of women and men*.
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, 54(6), 408–423. <https://doi.org/10.1037/0003-066X.54.6.408>
- Eagly, A. H., & Wood, W. (2012). Social Role Theory. In P. Van Lange, A. Kruglanski, & E. Higgins, *Handbook of Theories of Social Psychology* (pp. 458–476). SAGE Publications Ltd. <https://doi.org/10.4135/9781446249222.n49>
- Eagly, A., & Hall, J. A. (2024). *How Accurate are Gender Stereotypes? Evidence from Multiple Criteria and Multiple Methods*. <https://doi.org/10.2139/ssrn.4935106>
- Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In *Achievement and achievement motives* (pp. 75–146). Freeman.
- Eccles, J. S., & Wang, M.-T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, 40(2), 100–106. <https://doi.org/10.1177/0165025415616201>
- Else-Quest, N. M., & Grabe, S. (2012). The Political Is Personal: Measurement and Application of Nation-Level Indicators of Gender Equity in Psychological Research. *Psychology of Women Quarterly*, 36(2), 131–144. <https://doi.org/10.1177/0361684312441592>
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127. <https://doi.org/10.1037/a0018053>
- Eriksson, K., Björnstjerna, M., & Vartanova, I. (2020). The Relation Between Gender Egalitarian Values and Gender Differences in Academic Achievement. *Frontiers in Psychology*, 11, 236. <https://doi.org/10.3389/fpsyg.2020.00236>
- Esposito, L., Tonizzi, I., Usai, M. C., & Giofrè, D. (2025). Understanding the Role of Cognitive Abilities and Math Anxiety in Adolescent Math Achievement. *Journal of Intelligence* 13: 44. <https://doi.org/10.3390/>
- Fellman, D., Bränström, R., & Herlitz, A. (2021). Revisiting a basic question: Does growing up in either female or male environment affect sex differences in academic strengths and occupational choices? *Humanities and Social Sciences Communications*, 8(1), 221. <https://doi.org/10.1057/s41599-021-00897-5>
- Fors Connolly, F., Goossen, M., & Hjerem, M. (2020). Does Gender Equality Cause Gender Differences in Values? Reassessing the Gender-Equality-Personality Paradox. *Sex Roles*, 83(1–2), 101–113. <https://doi.org/10.1007/s11199-019-01097-x>

- Freiberg, T. (2019). Effects of Care Leave and Family Social Policy: Spotlight on the United States. *American Journal of Economics and Sociology*, 78(4), 1009–1037. <https://doi.org/10.1111/ajes.12293>
- Friedman, R. C., & Downey, J. I. (2002). *Sexual orientation and psychoanalysis* (pp. 225-263).
- Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74(2), 435–452. <https://doi.org/10.1037/0022-3514.74.2.435>
- Fryer, R. G., & Levitt, S. D. (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, 2(2), 210–240. <https://doi.org/10.1257/app.2.2.210>
- Ganley, C. M., & Vasilyeva, M. (2011). Sex differences in the relation between math performance, spatial skills, and attitudes. *Journal of Applied Developmental Psychology*, 32(4), 235–242. <https://doi.org/10.1016/j.appdev.2011.04.001>
- Geary, D. C. (2010). *Male, female: The evolution of human sex differences* (2nd ed.). American Psychological Association. <https://doi.org/10.1037/12072-000>
- Geary, D. C. (2015). *Evolution of vulnerability: Implications for sex differences in health and development*. Academic Press.
- Geary, D. C. (2024). The Evolved Mind and Modern Education: Status of Evolutionary Educational Psychology. *Elements in Applied Evolutionary Science*. <https://doi.org/10.1017/9781009454858>
- Geary, D. C., Hoard, M. K., Nugent, L., Chu, F., Scofield, J. E., & Ferguson Hibbard, D. (2019). Sex differences in mathematics anxiety and attitudes: Concurrent and longitudinal relations to mathematical competence. *Journal of Educational Psychology*, 111(8), 1447–1461. <https://doi.org/10.1037/edu0000355>
- Gevrek, Z. E., Gevrek, D., & Neumeier, C. (2020). Explaining the gender gaps in mathematics achievement and attitudes: The role of societal gender equality. *Economics of Education Review*, 76, 101978. <https://doi.org/10.1016/j.econedurev.2020.101978>
- Gevrek, Z. E., Neumeier, C., & Gevrek, D. (2018). Explaining the gender test score gap in mathematics: The role of gender inequality. *Available at SSRN 3111114*.
- Ghasemi, E., Burley, H., & Safadel, P. (2019). Gender Differences in General Achievement in Mathematics: An International Study. *New Waves-Educational Research and Development Journal*, 22(1), 27–54.
- Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science*, 320(5880), 1164–1165.
- Halpern, D. F. (2013). *Sex differences in cognitive abilities*. Psychology press.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The Science of Sex Differences in Science and Mathematics. *Psychological Science in the Public Interest*, 8(1), 1–51. <https://doi.org/10.1111/j.1529-1006.2007.00032.x>
- Halsey, L. G., Giofrè, D., & Geary, D. C. (2024). Does greater variation reside in the larger sex? *Biology Letters*, 20(11), 20240404. <https://doi.org/10.1098/rsbl.2024.0404>
- Herlitz, A., Hönig, I., Hedebrant, K., & Asperholm, M. (2024). A Systematic Review and New Analyses of the Gender-Equality Paradox. *Perspectives on Psychological Science*, 17456916231202685. <https://doi.org/10.1177/17456916231202685>
- Hervías Parejo, V., & Radulović, B. (2023). Public Policies on Gender Equality. In D. Vujadinović, M. Fröhlich, & T. Giegerich (Eds.), *Gender-Competent Legal Education* (pp. 405–428). Springer International Publishing. https://doi.org/10.1007/978-3-031-14360-1_12
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Psychological Assessment Resources.
- Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, 78(2), 250–261. <https://doi.org/10.1037/0021-9010.78.2.250>

- Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences*, *106*(22), 8801–8807. <https://doi.org/10.1073/pnas.0901265106>
- Ireson, G. (2017). Gender achievement and social, political and economic equality: A European perspective. *Educational Studies*, *43*(1), 40–50. <https://doi.org/10.1080/03055698.2016.1237868>
- Jonsson, J. O. (1999). Explaining Sex Differences in Educational Choice An Empirical Assessment of a Rational Choice Model. *European Sociological Review*, *15*(4), 391–404. <https://doi.org/10.1093/oxfordjournals.esr.a018272>
- Jussim, L., Crawford, J. T., & Rubinstein, R. S. (2015). Stereotype (In)Accuracy in Perceptions of Groups and Individuals. *Current Directions in Psychological Science*, *24*(6), 490–497. <https://doi.org/10.1177/0963721415605257>
- Kahn, S., & Ginther, D. (2017). *Women and STEM* (w23525; p. w23525). National Bureau of Economic Research. <https://doi.org/10.3386/w23525>
- Kaiser, T. (2019). Nature and evoked culture: Sex differences in personality are uniquely correlated with ecological stress. *Personality and Individual Differences*, *148*, 67–72. <https://doi.org/10.1016/j.paid.2019.05.011>
- Kane, J. M., & Mertz, J. E. (2012). Debunking Myths about Gender and Mathematics Performance. *Notices of the American Mathematical Society*, *59*(01), 10. <https://doi.org/10.1090/noti790>
- Kuhn, A., & Wolter, S. C. (2022). Things versus People: Gender differences in vocational interests and in occupational preferences. *Journal of Economic Behavior & Organization*, *203*, 210–234.
- Lee, K., & Ashton, M. C. (2020). Sex differences in HEXACO personality characteristics across countries and ethnicities. *Journal of Personality*, *88*(6), 1075–1090. <https://doi.org/10.1111/jopy.12551>
- Levanon, A., & Grusky, D. B. (2018). Why Is There Still So Much Gender Segregation? In *Inequality in the 21st Century* (pp. 371–379). Routledge.
- Liou, P.-Y., Lin, Y.-M., Huang, S.-C., & Chen, S. (2022). Gender Differences in Science Motivational Beliefs and Their Relations with Achievement over Grades 4 and 8: A Multinational Perspective. *International Journal of Science and Mathematics Education*. <https://doi.org/10.1007/s10763-021-10243-5>
- Lippa, R. A. (2010a). Gender differences in personality and interests: When, where, and why? *Social and Personality Psychology Compass*, *4*(11), 1098–1110.
- Lippa, R. A. (2010b). Sex Differences in Personality Traits and Gender-Related Occupational Preferences across 53 Nations: Testing Evolutionary and Social-Environmental Theories. *Archives of Sexual Behavior*, *39*(3), 619–636. <https://doi.org/10.1007/s10508-008-9380-7>
- Lofstedt, J. (2003). Gender and veterinary medicine. *The Canadian Veterinary Journal*, *44*(7), 533–535.
- Lynn, R., & Mikk, J. (2009). National IQs predict educational attainment in math, reading and science across 56 nations. *Intelligence*, *37*(3), 305–310. <https://doi.org/10.1016/j.intell.2009.01.002>
- Mac Giolla, E., & Kajonius, P. J. (2019). Sex differences in personality are larger in gender equal countries: Replicating and extending a surprising finding. *International Journal of Psychology*, *54*(6), 705–711. <https://doi.org/10.1002/ijop.12529>
- Martin, A. G. (2017). The New Egalitarianism. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2896397>
- Moakler, M. W., & Kim, M. M. (2014). College Major Choice in STEM: Revisiting Confidence and Demographic Factors. *The Career Development Quarterly*, *62*(2), 128–142. <https://doi.org/10.1002/j.2161-0045.2014.00075.x>
- Moè, A., Hausmann, M., & Hirnstein, M. (2021). Gender stereotypes and incremental beliefs in STEM and non-STEM students in three countries: Relationships with performance in cognitive tasks. *Psychological Research*, *85*(2), 554–567.

- Murphy, S. A., Fisher, P. A., & Robie, C. (2021). International comparison of gender differences in the five-factor model of personality: An investigation across 105 countries. *Journal of Research in Personality, 90*, 104047. <https://doi.org/10.1016/j.jrp.2020.104047>
- Napp, C., & Breda, T. (2022). The stereotype that girls lack talent: A worldwide investigation. *Science Advances, 8*(10), eabm3689. <https://doi.org/10.1126/sciadv.abm3689>
- OECD. (2009). *PISA Data Analysis Manual: SPSS, Second Edition*. OECD. <https://doi.org/10.1787/9789264056275-en>
- Ott-Holland, C. J., Huang, J. L., Ryan, A. M., Elizondo, F., & Wadlington, P. L. (2013). Culture and vocational interests: The moderating role of collectivism and gender egalitarianism. *Journal of Counseling Psychology, 60*(4), 569–581. <https://doi.org/10.1037/a0033587>
- Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting Intellectual Patterns Predict Creativity in the Arts and Sciences: Tracking Intellectually Precocious Youth Over 25 Years. *Psychological Science, 18*(11), 948–952. <https://doi.org/10.1111/j.1467-9280.2007.02007.x>
- Pelley, E., & Carnes, M. (2020). When a Specialty Becomes “Women’s Work”: Trends in and Implications of Specialty Gender Segregation in Medicine. *Academic Medicine, 95*(10), 1499–1506. <https://doi.org/10.1097/ACM.0000000000003555>
- Penner, A. M., & CadwalladerOlsker, T. (2012). (Hyde and Mertz, 2009). In *Stoet and Geary*, (pp. 441–468). Springer.
- Prix, I., & Kilpi-Jakonen, E. (2022). Not in a Class of One’s Own: Social Origin Differentials in Applying to Gender-(A)Typical Fields of Study across the Educational Hierarchy. *European Sociological Review, 38*(6), 920–941. <https://doi.org/10.1093/esr/jcac007>
- Rawlings, C. M. (2007). *Higher Education as a Prism: The Role of Organizational Structures in the Gender Segregation and Stratification of American Undergraduates, 1970–1995*. Working paper. University of California, Santa Barbara, Department of Sociology.
- Reilly, D. (2012). Gender, Culture, and Sex-Typed Cognitive Abilities. *PLoS ONE, 7*(7), e39904. <https://doi.org/10.1371/journal.pone.0039904>
- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education, 49*(1), 25–50. <https://doi.org/10.1007/s11165-017-9630-6>
- Richardson, S. S., Reiches, M. W., Bruch, J., Boulicault, M., Noll, N. E., & Shattuck-Heidorn, H. (2020). Is There a Gender-Equality Paradox in Science, Technology, Engineering, and Math (STEM)? Commentary on the Study by Stoet and Geary (2018). *Psychological Science, 31*(3), 338–341. <https://doi.org/10.1177/0956797619872762>
- Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The More Things Change, the More They Stay the Same? Prior Achievement Fails to Explain Gender Inequality in Entry Into STEM College Majors Over Time. *American Educational Research Journal, 49*(6), 1048–1073. <https://doi.org/10.3102/0002831211435229>
- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*(5), 667–706. <https://doi.org/10.1002/per.634>
- Sax, L. J., Kanny, M. A., Riggers-Piehl, T. A., Whang, H., & Paulson, L. N. (2015). “But I’m Not Good at Math”: The Changing Salience of Mathematical Self-Concept in Shaping Women’s and Men’s STEM Aspirations. *Research in Higher Education, 56*(8), 813–842. <https://doi.org/10.1007/s11162-015-9375-x>
- Schmitt, D. P., Long, A. E., McPhearson, A., O’Brien, K., Remmert, B., & Shah, S. H. (2017). Personality and gender differences in global perspective. *International Journal of Psychology, 52*, 45–56. <https://doi.org/10.1002/ijop.12265>
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can’t a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*(1), 168–182. <https://doi.org/10.1037/0022-3514.94.1.168>

- Schwartz, S. H. (2007). Basic human values: Theory, measurement, and applications. *Revue Française de Sociologie*, 47(4), 929.
- Schwartz, S. H., & Rubel-Lifschitz, T. (2009). Cross-national variation in the size of sex differences in values: Effects of gender equality. *Journal of Personality and Social Psychology*, 97(1), 171.
- Sohn, K. (2015). The influence of birth season on height: Evidence from Indonesia: The Influence of Birth Season on Height. *American Journal of Physical Anthropology*, 157(4), 659–665. <https://doi.org/10.1002/ajpa.22763>
- Stearns, E., Bottia, M. C., Giersch, J., Mickelson, R. A., Moller, S., Jha, N., & Dancy, M. (2020). Do Relative Advantages in STEM Grades Explain the Gender Gap in Selection of a STEM Major in College? A Multimethod Answer. *American Educational Research Journal*, 57(1), 218–257. <https://doi.org/10.3102/0002831219853533>
- Stewart-Williams, S., & Halsey, L. G. (2021). Men, women and STEM: Why the differences and what should be done? *European Journal of Personality*, 35(1), 3–39. <https://doi.org/10.1177/0890207020962326>
- Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with Higher Levels of Gender Equality Show Larger National Sex Differences in Mathematics Anxiety and Relatively Lower Parental Mathematics Valuation for Girls. *PLOS ONE*, 11(4), e0153857. <https://doi.org/10.1371/journal.pone.0153857>
- Stoet, G., & Geary, D. C. (2013). Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data. *PLoS ONE*, 8(3), e57988. <https://doi.org/10.1371/journal.pone.0057988>
- Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, 48, 137–151. <https://doi.org/10.1016/j.intell.2014.11.006>
- Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581–593. <https://doi.org/10.1177/0956797617741719>
- Stoet, G., & Geary, D. C. (2019). A simplified approach to measuring national gender inequality. *PLOS ONE*, 14(1), e0205349. <https://doi.org/10.1371/journal.pone.0205349>
- Stoet, G., & Geary, D. C. (2020). Sex-specific academic ability and attitude patterns in students across developed countries. *Intelligence*, 81, 101453. <https://doi.org/10.1016/j.intell.2020.101453>
- Stoet, G., & Geary, D. C. (2022). Sex differences in adolescents' occupational aspirations: Variations across time and place. *PLOS ONE*, 17(1), e0261438. <https://doi.org/10.1371/journal.pone.0261438>
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, 135(6), 859–884. <https://doi.org/10.1037/a0017364>
- Tao, C., Glosenberg, A., Tracey, T. J. G., Blustein, D. L., & Foster, L. L. (2022). Are Gender Differences in Vocational Interests Universal?: Moderating Effects of Cultural Dimensions. *Sex Roles*, 87(5–6), 327–349. <https://doi.org/10.1007/s11199-022-01318-w>
- Uunk, W. (2024). The Role of Math and Language Performance in Explaining the Gender Gap in STEM Major Choice. A Test for Germany. *Zeitschrift Für Soziologie*, 53(4), 404–418. <https://doi.org/10.1515/zfsoz-2024-2025>
- van de Werfhorst, H. G. (2017). Gender Segregation across Fields of Study in Post-Secondary Education: Trends and Social Differentials. *European Sociological Review*, 33(3), 449–464. <https://doi.org/10.1093/esr/jcx040>
- van der Slik, F. W. P., van Hout, R. W. N. M., & Schepens, J. J. (2015). The Gender Gap in Second Language Acquisition: Gender Differences in the Acquisition of Dutch among Immigrants from 88 Countries with 49 Mother Tongues. *PLOS ONE*, 10(11), e0142056. <https://doi.org/10.1371/journal.pone.0142056>

- Vianello, M., Schnabel, K., Sriram, N., & Nosek, B. (2013). Gender differences in implicit and explicit personality traits. *Personality and Individual Differences, 55*(8), 994–999. <https://doi.org/10.1016/j.paid.2013.08.008>
- Vishkin, A. (2022). Queen’s Gambit Declined: The Gender-Equality Paradox in Chess Participation Across 160 Countries. *Psychological Science, 33*(2), 276–284. <https://doi.org/10.1177/09567976211034806>
- Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence, 38*(4), 412–423. <https://doi.org/10.1016/j.intell.2010.04.006>
- Wai, J., Hodges, J., & Makel, M. C. (2018). Sex differences in ability tilt in the right tail of cognitive abilities: A 35-year examination. *Intelligence, 67*, 76–83. <https://doi.org/10.1016/j.intell.2018.02.003>
- Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology, 102*(4), 860–871. <https://doi.org/10.1037/a0019454>
- Wan, S., Lauermaun, F., Bailey, D. H., & Eccles, J. S. (2023). Girls’ comparative advantage in language arts explains little of the gender gap in math-related fields: A replication and extension. *Proceedings of the National Academy of Sciences, 120*(40), e2305629120. <https://doi.org/10.1073/pnas.2305629120>
- Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*(4), 304–340. <https://doi.org/10.1016/j.dr.2013.08.001>
- Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review, 29*(1), 119–140. <https://doi.org/10.1007/s10648-015-9355-x>
- Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science, 24*(5), 770–775. <https://doi.org/10.1177/0956797612458937>
- Webb, R. M., Lubinski, D., & Benbow, C. P. (2002). Mathematically facile adolescents with math-science aspirations: New perspectives on their educational and vocational development. *Journal of Educational Psychology, 94*(4), 785–794. <https://doi.org/10.1037/0022-0663.94.4.785>
- West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society, 1*(2), 125–151.
- West, C., & Zimmerman, D. H. (2009). Accounting for Doing Gender. *Gender & Society, 23*(1), 112–122. <https://doi.org/10.1177/0891243208326529>
- Winston, E. (1935). An Experiment with High-School Senior Vocational Interviews. *Journal of Educational Sociology, 9*(4), 195. <https://doi.org/10.2307/2262510>
- Wittmann, W. W. (2005). Group Differences in Intelligence and Related Measures. In O. Wilhelm & R. Engle, *Handbook of Understanding and Measuring Intelligence* (pp. 223–240). SAGE Publications, Inc. <https://doi.org/10.4135/9781452233529.n13>
- Wood, W., & Eagly, A. H. (2013). Biology or Culture Alone Cannot Account for Human Sex Differences and Similarities. *Psychological Inquiry, 24*(3), 241–247. <https://doi.org/10.1080/1047840X.2013.815034>
- Zhou, Y., Fan, X., Wei, X., & Tai, R. H. (2017). Gender Gap Among High Achievers in Math and Implications for STEM Pipeline. *The Asia-Pacific Education Researcher, 26*(5), 259–269. <https://doi.org/10.1007/s40299-017-0346-1>



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

ISBN 978-952-02-0215-6 (PRINT)
ISBN 978-952-02-0216-3 (PDF)
ISSN 0082-6987 (Print)
ISSN 2343-3191 (Online)