Academic Resilience in International Education Studies

Validity Challenges and Methodological Considerations

Wangqiong Ye

Dissertation submitted for the degree of PhD

Centre of Educational Measurement
Faculty of Educational Sciences

UNIVERSITY OF OSLO

2023

Series of dissertations submitted to the
Faculty of Educational Sciences, University of Oslo
No. 375

ISSN 1501-8962

Cover: UiO.
Print production: Graphic center, University of Oslo.

Acknowledgments

The rain brings forth an array of blessings, including the delights of Longjing tea, the juiciness of Yangmei berries, the unexpected break of typhoon holidays, and the fulfilling journey of my Ph.D. life.

The kickoff of my Ph.D. adventure started on a gently drizzling day in Dortmund, which proved to be typical weather for the locale, as evidenced by a collection of mushroom photos stored on my cellphone. As I stepped out of the U-Bahn station, I saw Sigrid Blömeke and Rolf Vegar Olsen standing there, waiting for me. The first meeting with my supervisors was brief, and right after, they sent me to the classroom, where Leonidas Kyriakides was giving a lecture about educational effectiveness research and dynamic models. To my delight, I found the rest of the OCCAM gang, all 14 of them, sitting there with beaming smiles. Little did I know then that we would go on to share countless moments and create cherished memories throughout the years to come.

Following the workshop in Dortmund, I set off on my adventure toward Oslo. As Johan Braeken's data science class started, my journey in CEMO began to gain momentum on its designated path. Indeed, there was quite a bit of chaos at the outset, but I found myself in capable hands and received abundant assistance. Rolf proactively reached out to the UDI to expedite my visa processing, contacted the educational faculty to secure my enrollment in the research design course just two days before its commencement, and provided invaluable guidance in developing my thesis. Björn Andersson kindly shared course materials with me after each of his classes, considering my lack of access to the Canvas platform at the time. Johan posed critical questions concerning the measurement of academic resilience in my study, and the invaluable insights gained from Brown Bag Seminars and Research Seminars contributed significantly to the evolution of my thesis. I am grateful to Stefan Schauber and Ronny Scherer for lending their ears and patiently listening to me, even though it required considerable time. Sigrid, serving as my main supervisor, played a pivotal role in my academic development, imparting knowledge and molding me into the researcher I am today. Not only did she provide unwavering support during moments of distress caused by journal rejections, but she also assisted me with extending my contract and offered valuable suggestions regarding future career prospects. In addition to the invaluable academic support, I received tremendous assistance from CEMO's dedicated administration team comprising Gunnhild Nedberg Grønlid, Tara Sarin, Amanda Marie Grini, and Anne Line Bjarke. Despite occasionally feeling like I burdened them with my concerns, they consistently demonstrated remarkable responsiveness, promptly addressing my inquiries, and effectively resolving any issues that arose. Apart from the work itself, life within CEMO is consistently delightful, as I am fortunate to be surrounded by a group of wonderful colleagues who possess diverse personalities, yet consistently exhibit kindness and camaraderie.

The experience in OCCAM has significantly enriched my Ph.D. studies, providing me with a valuable opportunity to engage with esteemed scholars at the forefront of the field. During my research stay in Dortmund, under the guidance of Rolf Strietholt, I completed Article 2 as part of my thesis. Additionally, Isa Steinmann and the international office of TU Dortmund played a pivotal role in facilitating my settlement and obtaining a residence permit, for which I am immensely grateful.

While I was working on the extended abstract for the thesis, Sigrid, unfortunately, fell ill and could not continue her supervision. Nevertheless, Nani Teig and Rolf stepped forward and assumed the responsibility of guiding and supervising me. The drafts of my thesis underwent several revisions, demanding a considerable amount of their time and energy. Trude Nilsen performed a thorough final reading of my work, providing me with detailed step-by-step instructions for revising the rationale section and highly detailed suggestions and feedback on the entirety of the thesis. Without their invaluable assistance, my thesis would not have progressed as smoothly and successfully as anticipated.

As I approach the completion of my thesis, my Ph.D. journey is drawing to a close. I am profoundly grateful for the presence and contributions of everyone who has been a part of this journey, as well as for the myriad experiences that have unfolded over the course of the past five years.

Summary

Academic resilience, usually focusing on disadvantaged students exhibiting favorable outcomes, has gained increasing attention because of its potential to mitigate disparities in academic performance and promote educational equity. This doctoral thesis explores the theoretical framework, methodological approaches, and empirical inquires related to academic resilience, with a specific focus on its relevance within the context of international education studies. The overarching aim of this thesis is two-fold. Firstly, it seeks to explore how academic resilience can be operationalized in order to function across the very different educational contexts represented in international large-scale assessments (ILSAs). Secondly, it aims to examine the methods that can be utilized to investigate protective factors at a global level. The empirical studies in this thesis are based on data from the Trends in International Mathematics and Science Study (TIMSS) 2019 and the Programme for International Student Assessment (PISA) 2015.

This thesis is based on three articles introduced and discussed in an extended abstract. The extended abstract includes a brief introduction to the theoretical framework, a discussion on validity challenges and methodological considerations related to academic resilience in international education studies, specifically focusing on the utilization of data from ILSAs. To bridge the research gaps outlined in the extended abstract, the three articles contribute to the overarching aim of the thesis by integrating theoretical, methodological, and empirical considerations in the study of academic resilience.

Article 1 endeavors to fill the research gap in the theoretical framework by conducting a comprehensive systematic review of academic resilience. Its primary objective is to summarize the impact of five distinct categories of protective factors (namely, individual, family, school, peer, and community) on the academic resilience of school-aged children and adolescents. The article explores the methodologies employed to investigate the relationships between these protective factors and academic resilience, while also examining the operationalization of academic resilience and the data utilized in these investigations. Article 1 identified research gaps pertaining to the operationalization and investigation of academic resilience. These identified gaps were further examined and addressed in Articles 2 and 3.

Article 2 directs its attention toward addressing the research gap concerning the measurement of academic resilience in the context of international studies. By employing four background indicators, namely a composite socio-economic status index and three indicators

representing economic, cultural, and social capitals within the family, as well as utilizing two distinct types of thresholds, Article 2 examines 16 different operationalizations of academic resilience. Within these 16 operationalizations, the article explores the component of resilient students and investigates the relationship between academic resilience and two external variables, with the aim of identifying a suitable definition within the specific context of international studies.

Article 3 is dedicated to addressing concerns surrounding the exploration of academic resilience in the context of international studies. It utilizes 11 protective factors associated with teacher quality, teaching quality, school resources, and school climate to discern latent profiles of resilient resources. Additionally, the article delves into examining the influence of educational expenditures on academic resilience within these four identified latent profiles, employing a three-step Bolck, Croon, and Hagenaars (BCH) method. This particular method, though infrequently employed within this field, holds promise as a methodological approach for studies on academic resilience.

The increasing utilization of ILSAs data in academic resilience research not only enhances comprehension of the impact of students' individual characteristics, family backgrounds, and learning environments, but also presents a valuable opportunity for conducting comparative studies across multiple countries. The findings presented in this doctoral thesis advance the existing knowledge about academic resilience, particularly its operationalization and research methods in international studies using ILSAs data. Furthermore, this study makes a noteworthy contribution towards informing researchers about the measurement and exploration of academic resilience, while also providing valuable insights for ILSAs and policymakers aiming to address achievement gaps and foster educational equity.

Table of contents

List of the articles

Article 1 **Ye, W.**, Teig, N., & Blömeke, S. (2023). Systematic review of protective factors related to academic resilience in children and adolescents: Unpacking the interplay of operationalization, data, and research method (under review in the journal of Educational Research Review)

Article 2 **Ye, W.**, Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. *Educational Assessment, Evaluation and Accountability, 33*(1), 169-202. https://doi.org/10.1007/s11092-020-09351-7

Article 3 **Ye, W.,** Olsen, R. V., & Blömeke, S. (2023). More money does not necessarily help: Relations of education expenditure, school characteristics, and academic resilience across 36 education systems (under review in the journal of Large-scale Assessments in Education)

List of the main abbreviations

ANOVA	Analysis of Variance
BCH	Bolck, Croon and Hagenaars Method
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
ESCS	Economic, Social, and Cultural Status
HER	Home Educational Resources
HRL	Home Resources for Learning
ILSA	International Large-Scale Assessment
IRT	Item Response Theory
LCA	Latent Class Analysis
LPA	Latent Profile Analysis
MANOVA	Multivariate Analysis of Variance
MG-CFA	Multiple-Group Confirmatory Factor Analysis
MG-SEM	Multiple-Group Structural Equation Modeling
MLM	Multilevel Modeling
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
RMSEA	Root Mean Square of Approximation
SEAS	School Emphasis on Academic Success
SEM	Structural Equation Modeling
SES	Socio-Economic Status
SRMR	Standardized Root Mean Square Residual
TIMSS	Trends in International Mathematics and Science Study

Part I Extended Abstract

1 Introduction

This doctoral thesis explores academic resilience within the context of international studies, using data from international large-scale assessments (ILSAs). The overarching objective of this research is twofold: firstly, to address validity challenges that arise from the operationalization of academic resilience in heterogeneous contexts; and secondly, to examine methodological considerations related to studying academic resilience in international studies. To achieve these objectives, data obtained from the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) are utilized.

The introduction chapter of this doctoral thesis begins with a contextual overview that frames the subsequent chapters. Following this, the chapter explains the principal goal of the thesis, elucidating the connection between the three included articles and the overarching aim. Lastly, the introduction provides a brief summary of all the chapters contained within the thesis.

1.1 Background and rationale

In recent decades, the concept of academic resilience has gained greater attention due to its capacity to ameliorate achievement gaps and foster educational equity. Broadly defined, academic resilience refers to the ability of students to perform well academically despite experiencing adversity. Investigations of academic resilience are closely related to three key concepts: risk, positive adaptation, and protective factor. Risk and positive adaptation constitute two fundamental dimensions for assessing academic resilience. The former refers to individual (i.e., ethnicity) or social factors (i.e., poverty) that are associated with a greater likelihood of poor development outcomes, while the latter refers to outcomes that surpass expectations in the presence of adversity (Tudor and Spray, 2017). On the other hand, protective factors refer to influences that modify, ameliorate, or alter a person's response to adversity (Rutter, 1987). Exploring the influence of protective factors on academic resilience has emerged as a critical focus in the field.

Research on academic resilience is urgently needed, not only to deepen our understanding of how some disadvantaged students manage to overcome adversities and excel academically, but also to shed light on how we might better support all students in achieving their academic potential. Such research can inform targeted interventions and education policies that are specifically tailored to address the distinct challenges encountered by students and more efficient allocation of resources, thereby contributing to equity in education. Research on academic resilience is not just a matter of academic interest but also a matter of urgency to promote equity and social justice in education. The persistent performance gap underscores glaring inequalities in educational outcomes, which run counter to the principles of fairness and social justice. By illuminating the factors that enable disadvantaged students to succeed academically, this research seeks to contribute to the development of educational practices and policies that can ensure all children, irrespective of their background, have a fair chance to reach their academic potential.

Due to its strong foundation in resilience studies within the domains of psychology and sociology, research on academic resilience has manifested in two distinct approaches, characterized by divergent theoretical frameworks. The first approach considers resilience an innate personal attribute, while the second conceptualizes it as a dynamic interplay between the individual and their environment. Consequently, the conceptualization and operationalization of academic resilience have diverged between the two approaches, leading to differences in the factors examined, research methods employed, and data utilized. For example, the approach that regards academic resilience as an inherent personal characteristic involves measuring it through a series of items or scales, which assess a student's ability to "bounce back" in the face of adversity. In this framework, greater emphasis is placed on individual, psychological factors such as motivation, self-esteem, and sense of control, which are commonly scrutinized through statistical techniques such as confirmatory factor analysis (CFA) and structural equation modeling (SEM). Conversely, the approach that treats academic resilience as a dynamic process underscores the importance of the interplay between individual and contextual factors, such as those related to family, schools, peers, and the community. Given the hierarchical nature of the data, multilevel modeling (MLM) has emerged as a prevalent analytical tool for this approach.

Since the 2010s, there has been growing interest in investigating academic resilience utilizing ILSAs data, particularly in comparative studies across countries. Given the cross-sectional design of ILSAs data, academic resilience is commonly operationalized by employing a combination of two criteria, namely low socio-economic status (SES) and high academic performance. However, the construct of academic resilience was originally developed within a homogeneous setting, such as a particular country. Thus, when applied to a heterogeneous context that encompasses multiple countries with diverse backgrounds, scholars encounter numerous challenges related to two primary issues. The first relates to the need for country-specific considerations in operationalizing academic resilience. For instance, the classification of low-SES students in a highly developed country like Norway may not align with the categorization of low-SES students in a less developed country. The second concerns the applicability of existing research methods in the ILSA context. For instance, the challenges linked to testing measurement invariance may impede the implementation of latent constructs in the context of ILSAs, which in turn may hinder the application of statistical methods such as CFA and SEM. Similarly, the utilization of MLM in country-specific data may fail to yield statistically significant findings, owing to the consequent reduction in statistical power. This thesis makes a valuable contribution by offering a more precise identification and description of these issues, as well as exploring potential solutions to address these two overarching concerns.

1.2 The overarching aim

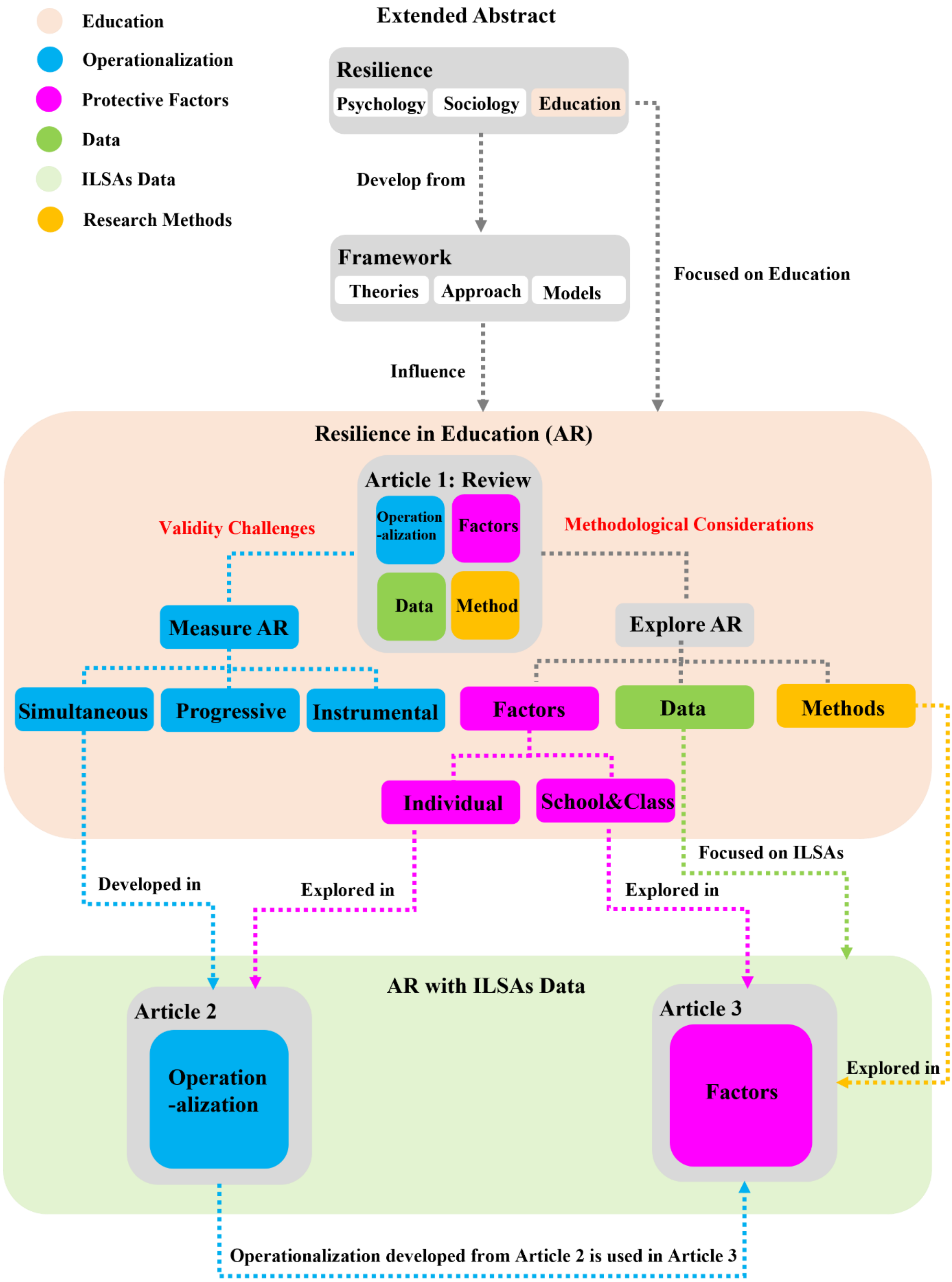
This doctoral thesis is designed to tackle the issues that arise when applying ILSA data in resilience studies, with two overarching aims. The first aim is to address issues related to the operationalization of academic resilience within international studies. The second aim is to explore analytical methods that can be employed to investigate the impact of protective factors on academic resilience across countries, utilizing ILSAs data.

Figure 1 provides an overview of the extended abstract and its connections to the three articles, which are visually represented by grey squares. The extended abstract begins with a brief introduction to the development of resilience in various fields. Subsequently, theories, research approaches, and analytical models that have been developed within these fields are briefly discussed, thereby establishing a foundation upon which the overarching aims of this thesis are built.

Next, the extended abstract directs its attention solely to academic resilience within the field of education, represented by the color “light peach” in Figure 1. The extended abstract delves deeper into examining the issues identified in Article 1, revolving around four key perspectives of academic resilience: operationalizations, data, protective factors, and research methods. These four perspectives are graphically depicted in Figure 1 using the colors blue, green, magenta, and orange, respectively. The extended abstract extends and delves deeper into the two primary issues examined in Article 1, namely validity challenges and methodological considerations. The former primarily focuses on the measurement of academic resilience, while the latter concentrates on the exploration of academic resilience and encompasses a broader spectrum of aspects, including protective factors, data considerations, and research methods. These identified issues are subsequently explored in greater depth within Articles 2 and 3.

Following that, Article 2 delves into the concern surrounding the measurement of academic resilience, while Article 3 focuses on the exploration of academic resilience, utilizing ILSAs data. These endeavors are visually represented by the color “light green” in Figure 1. In particular, Article 2 focuses on the first aim, namely, the operationalization of academic resilience in international studies. Article 3 primarily addresses the second aim by exploring research methodologies that can be employed to investigate protective factors that foster academic resilience in heterogeneous contexts.

Figure 1 *An Overview of the Extended Abstract and Three Articles*



Note. AR = Academic Resilience.

1.3 Overview of the articles

Article 1 presents a comprehensive systematic review of protective factors, operationalizations, data sources, and research methods utilized in resilience studies within the field of education. The challenges and issues identified in Article 1 were subsequently investigated further in the other two articles. Article 2 focuses on addressing operationalization issues, particularly those related to multiple dimensions of risk and thresholds for performance. Subsequently, the operationalization approach developed in Article 2 is applied in Article 3 to investigate the relationship between academic resilience and protective factors. To address issues concerning context considerations and statistical power, Article 3 utilizes latent profile analysis (LPA) to investigate the connections between education expenditure, protective factors within schools and classrooms, and academic resilience.

Article 1: Review

Ye, W., Teig, N., & Blömeke, S. (2023). Systematic review of protective factors related to academic resilience in children and adolescents: Unpacking the interplay of operationalization, data, and research method (under review in the journal of Educational Research Review)

This systematic review analyzed five distinct groups of *protective factors* (individual, family, school, peer, and community), in conjunction with three types of *operationalizations* for academic resilience (simultaneous, progressive, instrumental), three types of *data sources* (self-collected, national/local assessments, ILSAs), two timeframes (longitudinal and non-longitudinal), and commonly employed *research methods* in 119 empirical studies.

By establishing a linkage between protective factors, operationalization, data sources, and research methods, this article presented a systematic review of academic resilience from a comprehensive perspective. Furthermore, this article identified and examined the challenges and concerns that relate to these four perspectives in empirical studies.

Article 2: Operationalization

Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. *Educational Assessment, Evaluation and Accountability, 33*(1), 169-202. https://doi.org/10.1007/s11092-020-09351-7

The simultaneous operationalization identified in Article 1 was further examined in Article 2 to be applied in international studies utilizing ILSAs data. A combination of four *background indicators*, including a composite socio-economic status (SES) index, cultural, financial, and social capitals, along with two different *thresholds*, fixed and relative, were used to generate 16 operationalizations for academic resilience. These 16 operationalizations were evaluated in three distinct economies, namely Norway, Hong Kong, and Peru, utilizing PISA 2015 data. The primary objective of Article 2 was to investigate the degree to which the composition of disadvantaged students changes due to the application of various operationalizations and the extent to which the relationship between protective factors and academic resilience varies as a result of these different operationalizations.

Article 3: Protective Factors

Ye, W., Olsen, R. V., & Blömeke, S. (2023). More money does not necessarily help: Relations of education expenditure, school characteristics, and academic resilience across 36 education systems (under review in the journal of Large-scale Assessments in Education)

Aiming to address methodological considerations highlighted in Article 1, Article 3 focused on exploring the relationship between academic resilience and protective factors, including teacher quality, teaching quality, school resources, school climate, and education expenditure. The operationalization of academic resilience developed in Article 2 was applied.

This article aimed to identify *profiles of resilient resources* based on protective factors in schools and classrooms, utilizing latent profile analysis (LPA) and TIMSS 2019 data from 36 education systems. Furthermore, this study aimed to identify *cultural patterns* associated with these profiles within six distinct cultural groups, namely, Middle East, Post-Soviet, Confucian Asia, Anglo, Nordic, and Latin European countries. Additionally, utilizing a three-step Bolck, Croon, and Hagenaars (BCH) approach, this article sought to investigate the extent to which the association between education expenditure and academic resilience varies across these profiles.

1.4 Outline of the thesis

This thesis is divided into two parts: Part I, the extended abstract, and Part II, which includes three articles. To facilitate better comprehension, it is suggested that the articles in Part II be read before Part I, as they are referred to in the extended abstract. The six chapters in Part I provide a comprehensive background for the articles and a discussion on how the articles can be integrated and interpreted.

Chapter 1 provides an introduction to the motivation behind this thesis. The chapter briefly describes the background and rationale for this study, highlighting the challenges associated with exploring academic resilience in the context of international studies. To address these challenges, two overarching aims for the thesis are summarized for this thesis. This chapter also provides a brief overview of how each article in Part II contributes to these aims.

Chapter 2 provides an overview of the theoretical framework for resilience and academic resilience. Firstly, resilience is discussed in a broader context, including its development across various fields such as psychology, sociology, and education. The theories, approaches, and models developed in these fields are briefly summarized. Next, the discussion narrows down to academic resilience in education. An overview of operationalizations for academic resilience, data sources and timeframes, protective factors, and research methods is presented. Subsequent chapters will explore issues related to these four perspectives in greater detail.

Chapter 3 is dedicated to exploring the validity issues that arise from applying the concept of academic resilience, originally developed within a homogeneous context, to a heterogeneous context, such as international studies. This chapter addresses issues related to the measurement of risk, positive adaptations, and thresholds in the context of ILSAs. Specifically, it explores the multiple dimensions and missing data inherent in measuring risk, the domain differences and application of plausible values in identifying positive adaptations, and the challenges associated with setting appropriate thresholds for operationalizing academic resilience.

Chapter 4 focuses on methodological considerations, with a focus on three aspects. Firstly, the chapter examines challenges associated with ILSAs data, including differences in assessment design and small cluster sizes, as well as two primary approaches used in international studies. Secondly, statistical methods such as CFA, SEM, MLM, and latent class analysis (LCA) are discussed, and their application in the empirical studies of this thesis is described. Finally, the chapter addresses validity, reliability, and ethical considerations.

Chapter 5 provides a summary of the three articles presented in Part II.

In Chapter 6, the findings of these articles are further discussed to address the research gaps and overarching aim of this doctoral thesis. The contributions of this thesis to academic resilience are outlined, as well as its strengths and limitations. Finally, the chapter concludes with a brief remark.

2 Theoretical perspectives

2.1 Resilience

2.1.1 Historical roots and relevant fields

The concept of resilience, initially grounded in medicine, has been the subject of behavioral science research since the 1970s (Zolkoski & Bullock, 2012). In the mid-1980s, scholars from various disciplines, including child development, pediatrics, psychology, psychiatry, and sociology, investigated the phenomenon of resilience (Werner, 2000). Resilience research in education emerged around the 1990s, with roots in both psychology and sociology (Aburn, Gott, & Hoare, 2016).

The study of resilience encompasses a broad range of academic disciplines, with well-established fields frequently offering definitions and methodologies that are adopted and applied in other domains. Resilience research in psychology is highly predominant, typically adopting a longitudinal design to identify protective factors (Aburn et al., 2016). Resilience studies in sociology have an emphasis on the dynamic interaction between the individual and the environment, adopting a socio-ecological perspective (Ungar, 2011). Resilience studies in education frequently share a target population of children and adolescents with related research conducted in psychology and sociology (Aburn et al., 2016). Consequently, definitions and methodologies originating from psychology and sociology have been integrated, with requisite modifications and advancements made to accommodate the specific educational setting (Tudor & Spray, 2017).

2.1.2 The evolution of the resilience concept

Despite variations in context, the concept of resilience is closely related to an entity's capacity to return to a stable state after a disruption (Bhamra, Dani, & Burnard, 2011). Initially, resilience was conceptualized as individual traits or qualities that predict social and personal success. For example, Connor and Davidson (2003) defined resilience as "the personal qualities that enable one to thrive in the face of adversity." Inquiries examining resilience at this stage focus on identifying the characteristics of the individual at risk, for example, temperament, self-esteem, and planning skills (Buckner, Mezzacappa, & Beardslee, 2003; Hughes, Graham-Bermann, & Gruber, 2001). Later, with the development of the construct, the focus on resilience shifted to emphasize it as a dynamic process rather than personal trait (Fletcher & Sarkar, 2013). Luthar, Cicchetti, and Becker (2000) defined resilience as “a dynamic process encompassing positive adaptation within the context of significant adversity.” Resilience inquiries in this stage typically measure early risk and later positive adaptation, conceptualizing resilience as a capacity that develops over time during individual-context interactions (Fletcher & Sarkar, 2013). For example, Lansford et al. (2006) defined “physically abused during the first five years of life” as the risk and explored the associations between protective factors and children’s behavior from kindergarten through eighth grade. Although a universally accepted definition of resilience is lacking, it typically incorporates two core components: exposure to risk or adversity and the manifestation of positive adaptation (Luthar et al., 2000). Therefore, resilience is not directly measured but inferred based on directly measuring these two components (Luthar & Zelazo, 2003).

2.1.3 Categorizations for protective factors

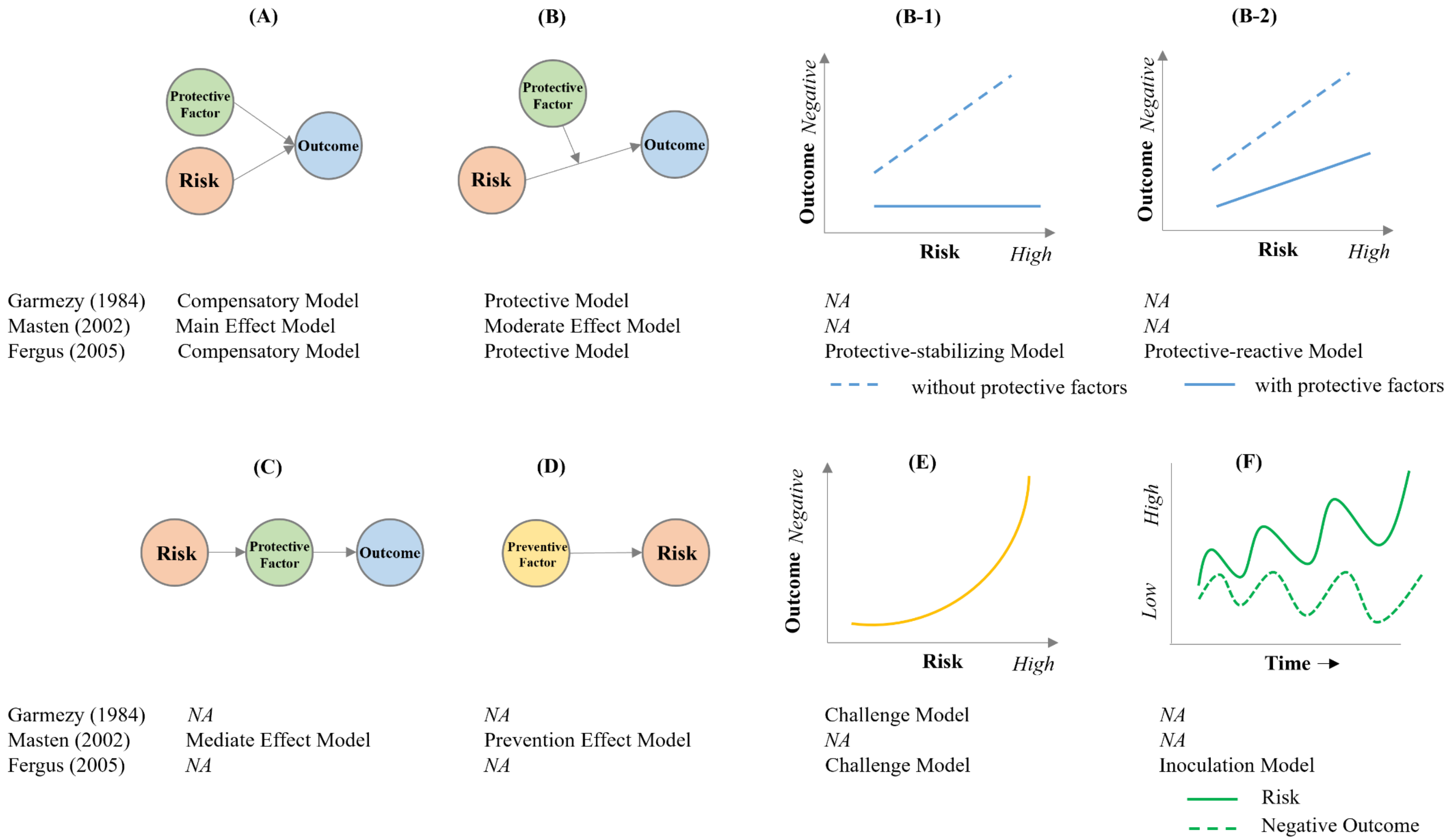
The conceptualization of resilience in research has varied based on adopting either a personal trait or dynamic progress definition. Nonetheless, both approaches have highlighted the presence of numerous protective factors related to resilience. Werner (2000) approached resilience from a personal trait perspective and categorized protective factors into distinct developmental periods: infancy (i.e., high levels of vigor, alertness, and sociability), childhood-adolescence (i.e., internal locus of control, strong achievement motivation, and positive self-concept), and adolescence-adulthood (i.e., planning, foresight). Later, with the shift towards a dynamic progress approach, greater attention was directed toward the interaction between the individual and the environment. Olsson et al. (2003) categorized protective factors into three distinct groups: individual, family, and social. Cutuli et al. (2016) undertook a similar classification and identified three protective factor categories, namely individual factors (i.e., positive outlook on life, self-regulation skills), family and close relationships (i.e., good parenting, organized home environment), community and connections with organizations (i.e., effective schools, neighborhood programs).

2.1.4 Resilience models and approaches

Researchers have developed various resilience models to explain how protective factors operate to alter the trajectory from risk exposure to positive adaptation. Garmezy, Masten, and Tellegen (1984) identified three resilience models: Compensatory (Figure 2A), Protective (Figure 2B), and Challenge (Figure 2E), which respectively proposed that protective factors have a direct effect on the outcome, moderate the impact of risk, or have a curvilinear relationship with risk. Fergus and Zimmerman (2005) extended the Protective model by proposing the Protective-stabilizing (Figure 2B-1) and Protective-reactive (Figure 2B-2) models and introducing the Inoculation model (Figure 2F). The Protective-stabilizing and Protective-reactive models differ in the degree to which they can neutralize the effects of risk. The former completely neutralizes the impact of risk and the latter only reduces the expected correlation between risk and outcome. The Inoculation model suggests that continuous exposure to low levels of risk can enhance resilience by preparing individuals to face more significant risks in the future.

Masten (2001) identified two approaches for resilience models: variable-focused and person-focused. The former approach employs multivariate statistics to examine the association between risk, protective factors, and positive adaptation, while the latter distinguishes resilient individuals from other groups by comparing individuals with different profiles based on specific criteria. Further, Masten and Reed (2002) found four resilience models for the variable-focused approach: Main effect, Moderate effect, Mediated effect, and Prevention effect. In the Main effect model, the protective factor contributes independently to the outcome, similar to the Compensatory Model in the study of Garmezy et al. (1984) (Figure 2A). The Moderate effect model proposes that protective factors moderate the relationship between risk and outcome, similar to the Protective Model in the study of Garmezy et al. (1984) (Figure 2B). Furthermore, Masten and Reed (2002) suggest two other models not previously captured in the literature: The Mediated Effect Model suggests that an influential factor on the outcome is impacted by both risks and protective factors (Figure 2C). The Prevention effect model, although uncommon in research, refers to the presence of a protective factor that prevents the occurrence of a risk factor (Figure 2D). For example, increasing adequate prenatal care can reduce prematurity rates, preventing the occurrence of the risk factor and its negative impact on later outcomes. For the person-focused approach, Masten and Reed (2002) specified three primary methods: 1) conducting single-case studies; 2) identifying a subgroup of high-risk individuals who exhibit resilience; and 3) utilizing pathway models that explain significant patterns in the life course.

Figure 2 *Resilience Models*



Note. Models were adapted from Garmezy, Masten, and Tellegen (1984), Masten and Reed (2002), and Fergus and Zimmerman (2005); *NA* = not discussed in the respective study.

2.2 Academic resilience

Although resilience studies in psychology and sociology provide the foundational definitions and methodologies, the integration into the domain of education has been carefully adapted to correspond with its distinctive attributes. In the context of education, the examination of resilience is centered on students, and consequently, the operationalizations and protective factors of resilience are primarily directed toward them. When applied to the field of education, the term academic resilience is often used.

Although several review papers have explored measuring academic resilience (Rudd, Meissel, & Meyer, 2021; Tudor & Spray, 2017), few studies have comprehensively reviewed the exploration of academic resilience. To address this research gap, Article 1 conducted a systematic review of academic resilience based on 119 empirical studies among school-aged students, focusing on protective factors, operationalizations, data, and research methods. This section provides a concise summary of Article 1, thereby presenting the most current overview of academic resilience.

2.2.1 Three operationalizations

The operationalization of academic resilience involves three primary approaches: 1) simultaneous approach, which measures both risk and positive adaptation at the same time; 2) progressive approach, which measures risk at an earlier time point and positive adaptation at a later time point, treating academic resilience as a developing process over time; and 3) instrumental approach, which measures academic resilience using multiple items or scales.

The simultaneous and progressive operationalizations include two core components, namely risk and positive adaptation (Tudor & Spray, 2017). Despite the shared focus on the interaction between the individual and context, the progressive operationalization prioritizes the dynamic nature of this interaction to a greater degree. Various factors are utilized to define risk, including low-SES, foster home residency, teenage motherhood, maltreatment, problematic parent-child relationships, and ethnic or minority status (see Appendix A). Conversely, positive adaptations are less diverse than risk factors, and primarily revolve around academic achievement. Examples of positive adaptations include not dropping out, returning to full-time education, completing high school, attending university, and adult attainment. In recent years, there has been a growing trend toward utilizing students' well-being to conceptualize positive adaptations (Austin et al., 2022). It is noteworthy that, unlike progressive operationalization, the simultaneous approach displays less variability in both risks and positive adaptations.

Instrumental operationalization differs significantly from the other two approaches. It usually includes items that describe cognitive-affective and behavioral responses to adversity, phrased positively or negatively. For example, Likert-scaled items such as “I would just give up” are used in the Academic Resilience Scale (ARS-30) developed by Cassidy (2016) to measure academic resilience. Because of diverse contexts and developers, various resilience scales may prioritize distinct aspects. For example, Skinner and Pitzer (2012) underscored the significance of students’ engagement and emotional reactivity, whereas Ungar and Liebenberg (2011) prioritized the role of peer and caregiver relationships.

2.2.2 Timeframes and data sources

Although simultaneous operationalization evaluates academic resilience at a single time point, the other two methodologies incorporate both longitudinal and non-longitudinal data. The progressive approach frequently employs longitudinal data to conceptualize risk and positive adaptation over time, whereas the instrumental approach typically utilizes longitudinal data to investigate the bidirectional association between protective factors and academic resilience.

In order to achieve their respective research objectives, scholars have utilized diverse datasets. Several studies investigating academic resilience have employed self-collected data, which provides researchers with greater control over the research process and allows for a more personalized approach. On the other hand, certain studies have acknowledged the drawbacks and potential biases of self-collected data, particularly stemming from low-response rates and challenges in achieving representative sampling. Consequently, some scholars have opted to employ pre-existing data sources such as national assessments and ILSAs.

The development of academic resilience, particularly in the United States during the early 2000s, was greatly influenced by national assessments. For example, Cappella and Weinstein (2001) and Wayman (2002) employed national assessment data and a progressive operationalization to investigate the interplay between academic resilience and school-level factors, including curriculum, school support, and the learning environment. Borman and Overman (2004), on the other hand, employed a simultaneous operationalization and investigated a comprehensive range of school-level protective factors, including school composition (percentages of low-achieving, minority, and free-lunch-eligible students), school resources (class size, instructional resources, and teachers’ experience), as well as leadership and teaching-related items (percent of classroom time spent on instruction and monitoring student progress). Moreover, they introduced a residual method to operationalize academic resilience by regressing students' SES on their academic performance and using the residual score for each student. This approach has since been widely adopted by numerous studies.

Several national assessments, such as the National Educational Longitudinal Study in the United States, are longitudinally designed and may be linked to data from government agencies or organizations, such as the Department of Human Services (Fantuzzo et al., 2012). Consequently, studies employing these assessments often operationalize academic resilience in a progressive approach, wherein risk is defined at an early stage and positive adaptation is assessed at a later time point.

Different from national assessments with longitudinal designs, ILSAs, such as PISA, often evaluate the performance of a specific age group of students every several years. Therefore, investigating the resilience of individual students over time is not feasible, as each cycle examines different cohorts of students. Despite the trend design of TIMSS and its administration every four years to grades 4 and 8 students, it does not track the same students over time, making it impossible to assess individual resilience across time. Thus, studies on resilience utilizing ILSAs data typically adopt a simultaneous operationalization to measure academic resilience. However, several scholars have employed ILSAs data to investigate academic resilience longitudinally on a national level. For instance, Agasisti and colleagues (2014a) calculated the percentage of resilient students at the country level and explored its association with education expenditure using five cycles of PISA data.

The utilization of ILSAs data in studies of resilience has increased since 2010, partly due to the evolution of these assessments. On the one hand, ILSAs data has significantly facilitated investigations of school-level protective factors, particularly teaching-related factors. On the other hand, ILSAs data has provided an opportunity to explore academic resilience across nations. Article 1 found that among the 19 empirical studies utilizing ILSAs data since 2010, 14 have conducted international comparisons of academic resilience. Due to the historical origins and subsequent evolution of the construct, academic resilience is frequently studied in a homogeneous context, such as a particular country, rather than a heterogeneous one. Consequently, additional challenges related to validity and methodological considerations arise when investigating academic resilience in an international context.

2.2.3 Protective factors

Among the five categories of protective factors, namely individual, family, school, peer, and community, the factors most commonly examined in education are those pertaining to individuals and schools. Individual factors have been extensively studied in the field of education, which may be attributed to their well-established prominence in resilience studies within psychology. In contrast, school-level factors have been more recently developed and explored in education.

The literature frequently examines four groups of individual factors: demographic characteristics, motivation and engagement, beliefs and attitudes, and learning practices or processes. Except for gender, demographic factors including race/ethnicity, immigration status/language, and age have generally exhibited a negative correlation with academic resilience, with minority status, immigration, and older age commonly associated with lower levels of resilience (Langenkamp, 2010; Wills & Hofmeyr, 2019). Empirical studies have consistently found that protective factors related to motivation and engagement are the strongest predictors of academic resilience, including students' motivation, aspirations, and engagement (Erberer et al., 2015; Garcia-Crespo et al., 2019). With the exception of self-esteem and attitude toward school, protective factors linked to beliefs and attitudes such as self-efficacy, confidence, sense of control, and valuing of school demonstrate a consistent positive association with academic resilience (Collie et al., 2017). Protective factors related to students' learning progress or practice have received less attention, with self-regulation being a consistent predictor of academic resilience (Koirikivi et al., 2021), whereas other factors, such as cognitive flexibility and work mastery, have produced inconsistent results (Süleyman, 2022).

Five groups of protective factors at the school level are frequently studied, including school material resources, discipline and climate, school academic enrichment and support, teacher quality, and teaching quality. The association between school material resources and academic resilience depends on the measurement instrument utilized. For instance, school average SES and location are consistently predictive of academic resilience (Agasisti et al., 2018), whereas school type, class size, and instructional resources yield mixed results (Vicente, Pastor, & Soler, 2021; Wills & Hofmeyr, 2019). Discipline and climate-related protective factors, such as discipline, a safe and orderly environment, and school emphasis on academic success, have consistently been found to predict academic resilience (Erberer et al., 2015; Gabrielli, Longobardi, & Strozza, 2022; Koirikivi et al., 2021). Academic resilience is positively associated with school academic enrichment and support factors, such as extracurricular activities and support during the transition (Agasisti et al., 2018). The relationship between teacher quality and academic resilience varies depending on the metric employed and the context in which it is investigated. For instance, while teachers with at least a bachelor's degree are associated with students' academic resilience in China, such a relationship does not hold in South Africa (Hofmeyr, 2019; Jin et al., 2022). The protective factors associated with teaching are primarily centered on the quality of teacher-student relationships and the teacher's instructional effectiveness. While the former has gained significant scholarly attention, the latter has received increased focus, especially in light of the emergence of ILSAs. The literature indicates a positive correlation between the teacher-student relationship and academic resilience (Bester & Kuyper, 2020), with teaching quality-related factors yielding inconclusive findings. Empirical investigations have confirmed the linkage between academic resilience with instructional clarity, classroom management, teachers' confidence in students, and their expectations for students (Bostwick et al., 2022).

Academic resilience research has devoted comparatively less attention to factors associated with family, peer, and community contexts. Regarding family-related factors, family resources have consistently been identified as predictors of academic resilience (Li & Yeung, 2019). However, other factors, such as family academic support and family structure and relationships, have yielded inconsistent relationships with academic resilience (Cheung et al., 2014; Cunningham & Swanson, 2010). Similarly, the association between academic resilience and peer relationships has produced inconclusive findings (Bellis et al., 2018), though peer support has been identified as a significant predictor of academic resilience (Koirikivi et al., 2021). The impact of community protective factors on academic resilience is contingent upon the specific measures employed; for instance, welfare has been found to predict academic resilience, whereas homeless shelters have not.

2.2.4 Research methods

Academic resilience studies employ both variable-focused and person-focused approaches, as identified by Masten and Reed (2002). However, the majority of investigations in this area have primarily explored the impact of protective factors on the relationship between risks and outcomes through the use of Moderate Effect (Figure 2B) and Mediate Effect (Figure 2C) Models. Other conceptual frameworks, such as Challenge (Figure 2E) or Inoculation (Figure 2F) Models, have received scant attention within the field (Fergus & Zimmerman, 2005).

While the person-focused approach is often used in qualitative research, such as interviewing a group of disadvantaged students, a combination of both approaches is typically utilized in quantitative studies. For instance, researchers may first identify a group of disadvantaged students using a person-focused approach, and then apply statistical analysis to investigate the influence of protective factors using a variable-focused approach.

Quantitative methods were the predominant research approach employed in the field. However, only a few empirical studies utilized mixed or qualitative methods, including case studies or interviews in combination with statistical analysis methods such as analysis of variance (ANOVA), correlation, and regression.

The conceptualization of academic resilience can vary in terms of the approach adopted, with different operationalizations treating it as a binary variable, a continuous variable, or a latent construct. Consequently, empirical research often employs statistical models such as logistic regression, linear regression, and structural equation modeling (SEM) to examine academic resilience. The incorporation of protective factors from multiple levels has generated a growing interest in multilevel modeling (MLM) for investigating academic resilience.

Although some scholars have chosen to disregard the hierarchical structure of data in favor of a more parsimonious model (Cheung, 2017; Li & Yeung, 2019), the majority of scholars recognize the importance of considering such structure in their analysis. In academic resilience studies utilizing MLM, a stepwise approach is commonly adopted (Agasisti et al., 2018). This methodology involves the fitting of a sequence of models, commencing with a baseline model that incorporates solely student-level variables. Subsequently, classroom or school-level factors are incrementally added to the models until a final model is achieved, which encompasses variables from all levels of analysis.

Nonetheless, when employing conventional methods such as SEM and the stepwise approach of MLM to analyze ILSAs data, a multitude of issues can emerge. Chapter 4 provides a detailed discussion of these issues. Additionally, conducting cross-national comparisons introduces further complexities for international studies, necessitating the consideration of country-specific characteristics in both research methods and the operationalization of academic resilience. These complexities will be explored in greater detail in the upcoming chapters.

3 Validity challenges in measuring academic resilience across countries

The concept of resilience in psychology and education was predominantly developed within a relatively homogeneous context, confined to specific countries. As a result, before the emergence of ILSAs, international studies on academic resilience were rarely addressed in the literature. However, some sociologists have conducted international studies on resilience and highlighted the moderating role of contextual and cultural factors. For example, protective processes may be valued and made available differently across diverse contexts and cultures (Ungar, 2008).

The advent of ILSAs facilitated international comparisons and consequently instigated inquiries into academic resilience on a global scale. Nonetheless, when researchers attempt to explore academic resilience across nations, they face an initial challenge related to the application of traditional concepts developed within a homogeneous context to a heterogeneous context.

ILSAs typically evaluate various student cohorts across cycles, leading researchers to adopt a simultaneous operationalization, where both risk and positive adaptation are assessed at the same time point. In resilience research utilizing ILSAs data, scholars commonly employ low-SES as an indicator of risk and high academic achievement as a measure of positive adaptation. Some scholars have extended the exploration of academic resilience beyond cognitive outcomes by incorporating student well-being as a defining criterion for positive adaptation (OECD, 2018). Despite these efforts, other essential aspects of construct validity are often neglected in empirical investigations.

Article 2 centered on tackling the challenges associated with operationalizing academic resilience in international studies. It mainly addressed issues concerning the dimensions of the composite SES index and thresholds utilized for measuring risk and performance. Considering the fact that the SES index in PISA covers a wide range of resources, Article 2 leveraged PISA data to investigate the impact of diverse dimensions of family capital. The discussion of thresholds in measuring academic resilience is notably scarce, as highlighted in Article 1. Nevertheless, as demonstrated in Article 2, the examination of threshold-related concerns becomes significantly salient within the realm of international studies. To account for the higher incidence of missing SES data among younger students, Article 3 opted to utilize TIMSS grade 8 data, which provided adequate information on Home Educational

Resources (HER) as an indicator of SES. Furthermore, all plausible values were utilized in the analyses performed in empirical Articles 2 and 3. The issues outlined in this chapter are primarily addressed in Article 2, with Article 3 also addressing them to some extent.

3.1 Measuring risks

3.1.1 Dimensions of composite SES indexes

Composite SES indexes are commonly utilized to identify disadvantaged students in resilience studies employing ILSAs data. Nonetheless, scant attention has been given to elucidating the specific dimensions of capital or resources captured by these indexes. For example, the composite SES index of PISA, Economic, Social and Cultural Status (ESCS), is based on data collected from 15-year-old students regarding their parents' occupation, parental educational attainment, and household possessions, including a combination of general and country-specific household items (OECD, 2019). In contrast, the composite SES index of TIMSS, Home Educational Resources (HER), is derived from responses provided by 8th-grade students regarding the number of books present in their home, the highest level of education attained by either parent, as well as the number of home study supports, such as internet connection and own room (Mullis & Martin, 2017).

The ESCS encompasses a broad range of capital or resources available to students, surpassing the scope of the HER scale. For instance, the ESCS incorporates a range of household possession items, such as the number of books, links to the internet, computers for school work, education software, televisions, and other items (please refer to Appendix B). On the other hand, the HER scale predominantly focuses on educational resources. When these composite SES indices are employed to identify disadvantaged students, the composition of the student population may exhibit variation. Consequently, the relationship between certain protective factors and academic resilience across studies using PISA and TIMSS may vary due to the difference in these SES indexes. An illustration is a study conducted by Chirkina et al. (2020), which examined the relationship between academic resilience and students' attitude toward mathematics, utilizing four items related to students' value, likeness, confidence, and engagement in mathematics. Their findings revealed stronger associations between academic resilience and students' attitude toward mathematics in the TIMSS dataset compared to the PISA dataset. This observation suggests that the composite SES index, which captures various aspects of family capital or resources, may influence the results obtained and consequently give rise to concerns regarding the validity of comparisons across studies.

It is worth noting that the composite SES index may not necessarily be based on identical items across countries. For example, each country was allowed to include up to three country-specific items in measuring students' household possessions, which is a component of the Economic, Social and Cultural Status (ESCS) index (Avvisati, 2020). Consequently, when utilizing these scales to measure risk, additional caution and scrutiny are warranted. Furthermore, the internal consistency of these composite SES indexes exhibits considerable variation between countries, which inherently impacts the comparability of corresponding effects across different nations (Rutkowski & Rutkowski, 2013).

3.1.2 Challenges of missing data

A number of ILSAs focus on younger student cohorts, such as fourth-grade students in TIMSS and the Progress in International Reading Literacy Study (PIRLS). In light of the potential limitations of younger students' capacity to comprehend, interpret and report family capital or resources, the composite SES index utilized in TIMSS and PIRLS, known as Home Resources for Learning (HRL), is derived from questionnaires completed by both students and their parents (Mullis & Martin, 2017; Mullis & Martin, 2019). The HRL scale encompasses two distinct items obtained from students, namely the number of books present in their household, and the number of home study supports available to them, such as internet connection and own room. In addition, three items are included from the parents, which relate to the highest level of education and occupation of either parent and the number of children's books in the home. The administrative complexities arising from discrepancies in response rates between students and parents result in missing information for HRL components. Consequently, it is not uncommon to observe a missing rate of approximately 20% or higher in the HRL scale in grade 4 data from TIMSS or PIRLS. As an illustration, New Zealand and Norway exhibit missing data rates of 59.83% and 43.28% on the HRL scale in TIMSS 2019, respectively (please refer to Appendix C). This poses significant challenges in accurately identifying disadvantaged students for studies on academic resilience.

3.2 Measuring positive adaptations

3.2.1 Domain differences

With rare exceptions, the majority of studies using ILSAs data have employed academic performance as a metric for assessing positive adaptations. Some scholars have utilized one specific domain such as mathematics (Cheung, 2017), while others have adopted scores from several domains, such as an average score in mathematics, science, and reading (Gabrielli et al., 2022). Many studies have examined resilience variations across domains. For example, OECD (2011) found that disadvantaged students who perform well in science usually also do well in mathematics and reading. When academic resilience was explored with protective factors, some scholars identified disparities between domains. As an illustration, Hofmeyr (2019) explored a group of protective factors and demographic characteristics in both the TIMSS and PIRLS datasets. The research findings indicate a significant association between gender and academic resilience in reading, but not in mathematics. Moreover, the enjoyment of reading is not a predictor of resilience in PIRLS, while the enjoyment of mathematics is a significant predictor in TIMSS. Similarly, Garcia-Crespo et al. (2019) investigated the influence of individual and classroom factors on academic resilience in mathematics and science. The results also revealed disparities between the two domains, with students' enjoyment for learning being a predictor of resilience in mathematics but not in science. In sum, the potential differences among domains are acknowledged by the majority of scholars in the field.

3.2.2 Plausible values

In contrast to the extensive attention given by scholars to the domain issue, there has been little attention given to the issue of plausible values. Due to time limits and the large number of test items contained in ILSAs, students usually received a booklet with a subset of test items (Von Davier, Gonzalez, & Mislevy, 2009). Therefore, a complex statistical methodology is used to estimate students' proficiency, which considers the measurement error and sampling error in the assessment results. To account for the measurement error, multiple plausible values are generated for each student based on their test scores and conditioned on their responses to the background variables. In other words, the plausible values represent the likely range of scores for students with similar background characteristics. Therefore, it is recommended to use all plausible values to accurately capture a student's true score (Laukaityte & Wiberg, 2017).

In resilience studies using ILSAs data, researchers frequently operationalize student performance through the employment of plausible values. However, the incorporation of all plausible values has not been consistently implemented in all investigations, albeit this practice has been more commonly observed in recent years (Jang, Seo, & Brutt-Griffler, 2023). Most studies establish a threshold for academic achievement and subsequently convert each plausible value into binary form (i.e., resilient = 1, non-resilient = 0). Nevertheless, some researchers have employed only one plausible value to distinguish high-achievers in practice, particularly when utilizing the residual method proposed by Borman and Overman (2004), which identifies resilient students through the regression analysis of SES and academic performance (Gabrielli et al., 2022; Jin et al., 2022). Certain researchers have identified high-achievers through the use of the mean score of plausible values (Martin et al., 2022; Özcan & Bulus, 2022), while others have classified high-achievers as students who have performed above a specific threshold in over half of the plausible values (Vicente et al., 2021).

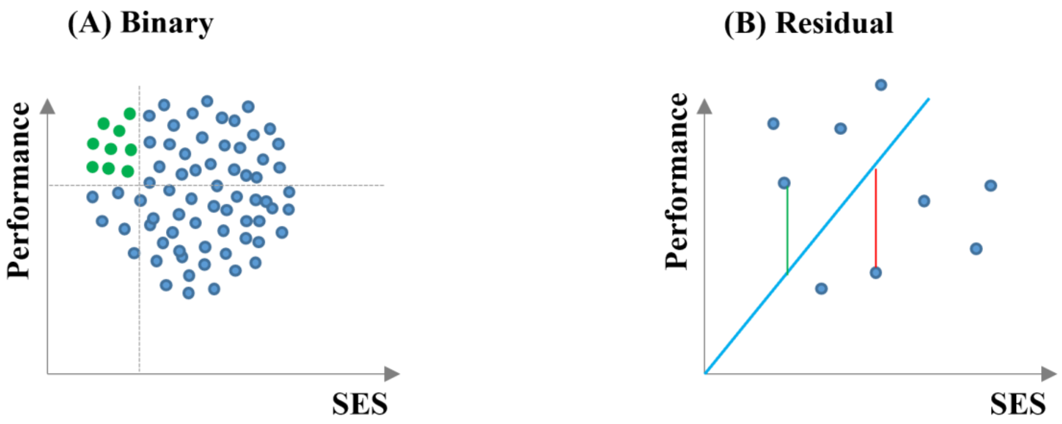
Moreover, an alternative type of resilience research aims to investigate the correlation between protective factors and the percentage of resilient students across countries (Agasisti et al., 2018). Typically, these investigations utilize a single plausible value to identify high-achieving students, and then calculate the percentage of resilience students. The significance of this research is rooted in the ability to investigate academic resilience longitudinally at the national level through the utilization of the percentage of resilient students. Thereby, this circumvents the constraints inherent in ILSAs, which assess different cohorts of students at each cycle. Nevertheless, a potential limitation of this approach is the likelihood of overlooking measurement errors.

The variation in the utilization of plausible values can be attributed in part to the use of different methodologies. For instance, the residual approach necessitates the inclusion of academic achievement, whereas the identification of high-achievers across different domains may encounter the challenge of utilizing several groups of plausible values. However, the potential measurement errors that arise from disregarding some plausible values are not extensively deliberated upon. Several scholars have argued that adopting the average of all plausible values constitutes a viable approach to exploring academic resilience, given that the measure of achievement is not utilized as the ultimate outcome but rather as a means to identify resilience status (Martin et al., 2022). Nevertheless, additional empirical investigations are required to explicate the impact of these methods.

3.3 Thresholds for risks and positive adaptations

Empirical studies using ILSAs data frequently employ two approaches, namely binary and residual, to identify resilient students. The binary method involves the direct application of two criteria: identifying disadvantaged students using a risk indicator (i.e., SES) and high-achievers using a performance indicator. Resilient students are recognized as those disadvantaged students who demonstrate high performance (see green dots in Figure 3A). Conversely, the residual method indirectly applies two criteria (Figure 3B). For instance, students' SES is regressed on their performance, and the residual obtained from the regression analysis is then allocated to individual students as their resilience scores. The empirical studies conducted within this thesis exclusively employed the binary approach, thus precluding further discussion on the residual approach within the scope of this research.

Figure 3 *Binary and Residual Approaches to Identify Resilient Students*



In the application of these two criteria, researchers often utilize two types of thresholds, namely fixed and relative, to define risk and positive adaptations. The fixed threshold involves the implementation of a consistent cut-off point across countries, while the relative threshold entails the adoption of different cut-off points that vary across countries.

3.3.1 Fixed and relative thresholds for risks

Considering the significant economic disparities among countries, employing a universal cut-off point may fail to capture accurate representations. Recognizing this, the utilization of a relative threshold is widely acknowledged as more appropriate when defining disadvantaged students on an international scale. This approach accounts for the unique socio-economic circumstances of each country and adjusts the defining criterion accordingly. By adopting a relative threshold, a more tailored definition of disadvantaged students can be established, acknowledging the diverse conditions present in different countries.

3.3.2 Fixed thresholds for performance

Different from the identification of disadvantaged students, the identification of high-achieving students has been observed through the application of both fixed and relative thresholds. Several scholars have employed a fixed threshold for academic performance, namely the same score universally applied across all nations, to evaluate academic resilience at a global level (Erberer et al., 2015; OECD, 2011).

Nonetheless, the application of a fixed threshold fails to consider the wide variations in average academic achievement across countries and may result in an inadequate understanding of what it means to be a high achiever in countries at the lower end of the achievement scale. For example, when a fixed threshold is applied to students’ performance, it is expected that high-achieving countries like Singapore may have a significant proportion of resilient students, reaching 80% or higher. In contrast, low-achieving nations such as Azerbaijan and Kyrgyzstan are anticipated to have a notably low proportion of resilient students, less than 5% (OECD, 2011). As a result, using fixed performance thresholds across countries with widely different average achievement scores leads to resilience being more or less defined by only this dimension of the resilience measure.

Scholars who advocate the use of fixed thresholds for academic performance contend that such an approach is designed to facilitate international comparisons (OECD, 2011). Nonetheless, the comparison being made is not actually between the academic performances of students worldwide. Rather, it is an evaluation of the degree to which students who are deemed as “disadvantaged” can overcome adversities and succeed academically. Consequently, it is advisable to take into account the local context when determining the benchmarks for academic success.

3.3.3 Relative thresholds for performance

In research employing a relative threshold on academic resilience, it is not uncommon to observe a slightly greater proportion of resilient students from some low-SES and low-achieving nations. For example, in Article 3, the application of a relative threshold to students’ performance revealed that the highest proportion of resilient students was not found in high-achieving countries like South Korea, but rather in South Africa, a country characterized by relatively lower levels of SES and academic performance. This phenomenon can be attributed in part to the comparatively lower difficulty in achieving academic success in such nations. However, it is also possible that the observed outcomes are influenced by the specific cut-off point utilized for the threshold, as well as the distribution of student performance around this point. For instance, if student performance is centered around the cut-off point, this could also impact the results.

On the one hand, the level of risk and academic success varies depending on the context. On the other hand, the degree to which the context enables a student to recover from adversity also varies across contexts. While the former can be assessed through indicators such as SES and academic performance, the latter is often implicit and not readily observable. Some scholars have proposed using the proportion of resilient students as a means of evaluating educational equity and quality (Agasisti et al., 2018). However, it is important to note that this metric mainly reflects the potential for disadvantaged students to overcome obstacles within a specific country, rather than as the sole indicator of that country's capacity to assist disadvantaged students. Therefore, greater circumspection is required when interpreting the findings regarding the presence of resilient students across the world.

3.4 Validity, reliability, and ethical considerations

3.4.1 Validity

The preceding sections have addressed the validity challenges pertaining to the measurement of academic resilience in international studies. The purpose of this particular section is to offer a more comprehensive discussion on the considerations of validity for academic resilience, along with other constructs examined within this thesis.

Validity refers to the extent to which a concept is accurately measured and encompasses several dimensions, including content, construct, and criterion validity (Heale & Twycross, 2015; Kane, 2006). Content validity refers to the extent to which a research instrument accurately measures all aspects of a construct (Heale & Twycross, 2015). The two aspects of content validity, relevance and representativeness, were taken into account in all empirical studies in this research (Yusoff, 2019). With regard to the construct of academic resilience, Article 1 summarized three commonly used operationalizations for defining this construct based on a systematic review of 119 studies. Article 2 narrowed the focus to studies utilizing ILSAs data and identified and discussed the most pertinent and representative criteria for determining academic resilience. In Article 2, 16 operationalizations for academic resilience were developed using four background indicators and two types of thresholds. Most of these operationalizations are widely employed in the field. Article 3 operationalized the construct of cognitive activation based on seven Likert-scaled items, which were answered by teachers on their classroom teaching practices. This approach is frequently used in the literature to measure cognitive activation, as noted by Blömeke, Olsen, and Suhl (2016).

Construct validity refers to the extent to which a research instrument measures the intended construct (Heale & Twycross, 2015). In this thesis, a confirmatory factor analysis (CFA) methodology was utilized to assess the construct validity. In Article 2, CFA was employed to assess the construct of students' sense of belonging to the school and their school attendance. Likewise, the construct of cognitive activation in Article 3 was evaluated using the CFA approach. All statistical analyses yielded a favorable fit of the models.

Criterion validity refers to the extent to which a research instrument is related to an outcome (Taherdoost, 2016). Regression analyses were employed in this research to measure the extent to which one measure predicts another. In Article 2, 16 operationalizations of academic resilience were assessed in conjunction with two external variables, namely, students’ sense of belonging to the school and their school attendance. Furthermore, the coefficients between these variables and academic resilience were compared across the 16 operationalizations, aiding in the identification of a more suitable definition for the construct in the realm of international education research.

Internal validity refers to the extent to which a study accurately demonstrates a causal relationship between variables, with a minimum of confounding variables or alternative explanations for the results (Flannelly, Flannelly, & Jankowski, 2018). Considering the cross-sectional design of ILSAs, this thesis did not address this validity. Instead, the research considered external validity, which refers to the degree to which inferences drawn from a given study’s sample apply to a broader population or other target populations (Findley, Kikuta, & Denly, 2021). Article 2 investigated the concept of academic resilience in Norway, Chile, and Hong Kong, which encompassed diverse academic achievements, cultural backgrounds, and economic development levels. In Article 3, the construct of cognitive activation was explored in 36 education systems, which included six cultural groups, namely, Middle East, Post-Soviet, Nordic, Anglo, Latin European, and Confucian Asia countries. By examining these constructs across diverse education systems and cultural groups, it is possible to make inferences about how the constructs are likely to manifest in other contexts.

Furthermore, this thesis utilized data from PISA and TIMSS studies, which assess 15-year-old and 8th-grade students, respectively, at the end of their compulsory education. Disadvantaged students were identified based on those who remained in school, yet it should be acknowledged that the most vulnerable students who dropped out may not have been included in this study. As such, the disadvantaged group identified in resilience studies may not be representative of the entire population. For example, while PISA 2018 covered around 90% of 15-year-olds in OECD countries, it only reached approximately 50% of this population in Jordan and Azerbaijan (OECD, 2019). It is crucial to acknowledge that the disadvantaged students identified in this study may not reflect the characteristics of the entire population, and thus, generalizing the findings to the broader population may be inappropriate. Therefore, policies and interventions developed based on these findings should be approached with caution, as they may not fully address the underlying issues related to educational equity and may neglect those who are most vulnerable.

3.4.2 Reliability

Reliability refers to the degree to which a measure consistently produces the same results over time or across different raters or contexts (Heale & Twycross, 2015). Internal consistency is a reliability measure that examines the consistency of responses to items used to form a scale (Yusoff, 2019). In this thesis, Cronbach's alpha was utilized to evaluate the internal consistency of most scales. Article 3 utilized Cronbach's alpha to evaluate the internal consistency of the cognitive activation construct, which yielded a high reliability score. Furthermore, Article 3 placed particular emphasis on multi-level raters. In this study, school-level variables rated by teachers were aggregated to the school level, while teaching-related variables rated by students were aggregated to the classroom level. This aggregation process effectively mitigated the potential measurement error originating from individual variations or random fluctuations in the responses of both teachers and students.

3.4.3 Ethical considerations

The current thesis employed ILSAs data from PISA and TIMSS studies, which are anonymous and publicly available, thus obviating the need for obtaining consent forms or approval from the Norwegian Centre for Research Data (NSD). Nonetheless, ethical concerns and challenges may arise when interpreting and applying the results derived from these data.

The central construct of this thesis, namely academic resilience, is intricately linked to the SES index incorporated in ILSAs such as PISA and TIMSS. However, as highlighted by Hopfenbeck and Kjærnsli (2016), students may perceive the questions in the PISA tests as overly intrusive regarding their family background, and some students may feel a sense of unease or even embarrassment regarding their parents' circumstances. Consequently, when students are required to rate such sensitive information, particularly items related to parental education, occupation, and home possession, several ethical considerations may arise, for example, potential emotional distress.

Similarly, students may feel uncomfortable evaluating the instructional quality of their teachers, while teachers themselves may also hesitate to assess school contributions or parental involvement in student academic success. Although participant confidentiality is upheld throughout the data collection process, and the sensitivity of these issues may not necessarily cause harm to the participants, it is crucial to maintain transparency regarding these potential ethical considerations.

4 Methodological considerations in exploring academic resilience across countries

Despite the increasing use of international studies with alternative data sources, it is worth noting that ILSAs data remains the dominant choice for such studies. These datasets typically cover numerous countries with diverse backgrounds and pose a range of methodological challenges that conventional methods, such as structural equation modeling (SEM) and the stepwise approach of multilevel modeling (MLM), do not usually address.

This chapter begins by discussing two significant issues related to the design of ILSAs and research approaches used in international studies. These issues were addressed in the doctoral thesis through empirical studies, which will be discussed in detail in the following sections. Subsequently, it examines issues related to several primary statistical methods employed in the field, including confirmatory factor analysis (CFA) and SEM, MLM, and latent class analysis (LCA). Articles 2 and 3 incorporated multilevel considerations, with Article 2 using CFA and SEM and Article 3 utilizing LCA to address specific research objectives. Moreover, Article 3 explored novel statistical methods to investigate protective factors within ILSAs data. Lastly, this chapter examines validity, reliability, and ethical considerations.

4.1 The use of ILSAs data

The ILSAs data from PISA, TIMSS, and PIRLS assessments are often utilized in exploring academic resilience across countries. Notwithstanding, due to distinctive assessment designs, targeted student populations, and domain variations inherent in these assessments, studies employing these data have demonstrated unique patterns in examining academic resilience.

4.1.1 PISA data

PISA utilizes a two-stage sampling technique, in which a minimum of 150 schools where 15-year-old students may be enrolled in each country are sampled at the first stage, and approximately 40 students[1] are randomly selected from a chosen sample school in the second

[1] In countries where paper-based assessments are employed, the typical number of students is 35, whereas in countries utilizing computer-based assessments, the number of students amounts to 42 in PISA 2018.

stage (OECD, 2019). Consequently, these students are not directly linked to their teachers or provided with classroom-specific information. A limited number of countries participating in PISA provide information about the learning environment, which is answered by teachers in the sampled schools[2]. However, it is important to note that the students and teachers in the same school sample may not be directly linked.

The empirical studies investigating protective factors, particularly those associated with teachers and schools, have demonstrated the sampling character of PISA. The investigation of academic resilience through PISA data has prioritized the examination of school-related factors, notably those of school resources, while classroom-level factors have received comparatively less attention.

Although the classroom is not considered a sampling unit in PISA, the study incorporates variables on teachers and teaching through its principal and student questionnaires. These questionnaires capture essential aspects such as the disciplinary climate within the classroom and the level of support provided by teachers, as reported by students. Additionally, principals provide information on teacher characteristics, including educational background and professional development (OECD, 2019). However, the absence of correspondence between teachers and students in PISA data has led to a tendency to analyze classroom-level factors (either aggregated from students' ratings or obtained from principals' ratings) at the school level, thereby overlooking the potential impact of individual classrooms on students' academic resilience. For example, teacher-related factors, including the percentage of qualified teachers within a school, the proportion of teachers with fixed-term contracts, and the mean number of years of teaching experience among the faculty, are commonly analyzed at the level of the school (Agasisti & Longobardi, 2014b; Cheung, 2017).

However, considering the fact that students spend most of their time in the classroom, the interaction between teachers and students plays a crucial role in promoting academic resilience, as evidenced by numerous studies (Bostwick et al., 2022; Garcia-Crespo, Fernandez-Alonso, & Muniz, 2021). Consequently, examining classroom factors solely at the school level overlooks a vital component of students' school experience. While a school-level perspective can offer valuable insights into administrative policies, resource allocation, and the overall climate of a school, adopting a classroom-level approach provides a more detailed and directly applicable understanding of everyday teaching and learning processes. Therefore,

[2] 19 participating economies completed the teacher questionnaire in PISA 2018.

the classroom level emerges as a particularly pertinent level of analysis when examining factors associated with academic resilience.

4.1.2 TIMSS and PIRLS data

Both TIMSS and PIRLS utilize a two-stage random sample design, wherein a sample of schools is selected as the first stage, followed by the selection of one or more intact classes of students from each of the selected schools as the second stage (Martin, Von Davier, & Mullis, 2020). Unlike the sample design in PISA, intact classes of students are sampled rather than individuals from across the grade level. This approach involves the nested grouping of students within their respective classrooms, alongside their teachers. This design enables researchers to examine the impact of teacher- and teaching-related factors on academic resilience. For instance, Erberer et al. (2015) employed TIMSS data to explore the influence of teachers' confidence and expectations on students' academic resilience. Similarly, Garcia-Crespo et al. (2021) analyzed a range of factors related to teacher quality and teaching quality, including teachers' basic and complementary training, job satisfaction, classroom instruction on reading strategies, and homework tracking, utilizing PIRLS data. Moreover, resilience studies utilizing TIMSS and PIRLS data have placed significant emphasis on school climate, including disciplinary climate, school emphasis on academic success, and a safe and orderly environment (Garcia-Crespo et al., 2019). Nevertheless, it is imperative to consider that within the context of TIMSS and PIRLS, multiple teachers specializing in a particular domain may have instructed the same classroom. Consequently, students' evaluations concerning their teachers' instructional methods may not necessarily pertain to the practices of a single teacher exclusively.

4.1.3 Cluster size of disadvantaged students

The issue related to cluster size has received limited attention within the existing literature. The PISA dataset for a given country typically comprises a sample of roughly 150 schools, each of which is represented by around 40 students (OECD, 2019). While the TIMSS and PIRLS datasets usually include a sample of approximately 4,000 students drawn from about 150 schools within a country (Martin et al., 2020). Specifically, there exist approximately 40 students per school in the PISA dataset, and around 25 students per classroom in the TIMSS or PIRLS dataset. When applying a bottom 1/3 threshold to the composite SES indicator to identify disadvantaged students, the cluster size for disadvantaged students is expected to be around 14 and 9 in the PISA and TIMSS or PIRLS datasets, respectively. Empirical studies often focus on disadvantaged students, and the analyses are thus conducted in this subset of at-risk students. Subsequently, the relatively small cluster size of disadvantaged students poses statistical power and reliability challenges in statistical analysis, particularly those focusing on a single country with limited cluster numbers.

The research objectives of empirical studies in this thesis were aligned with the ILSAs assessment designs. Article 2 employed PISA data, while Article 3 utilized TIMSS data to highlight classroom protective factors. Due to missing data issues concerning the composite SES index in PIRLS, its data were not included in this research. To address limitations arising from small cluster sizes of disadvantaged students, Article 3 employed latent profile analysis as an alternative to the commonly used multilevel modeling approach.

4.2 Analytical approaches in international studies

International studies that employ ILSAs data to examine the relationship between academic resilience and protective factors can be categorized into two groups: the first involves the application of pooled data to identify overarching trends, while the second involves the utilization of country-specific data to investigate variations between nations. Although the former approach yields more substantial outcomes across various studies due to a larger sample size, it is less context-sensitive. Conversely, the findings from the latter analytical approach are frequently challenging to generalize. Moreover, the outcomes from both approaches are prone to significant fluctuations contingent upon the variables and covariates integrated into the analysis, particularly those associated with MLM (Schoeneberger, 2016).

4.2.1 Aggregated trend analysis

The approach of aggregated trend analysis encompasses two distinct types of research. The first approach involves utilizing pooled data from a diverse set of countries to examine the correlation between protective factors and academic resilience (Agasisti & Longobardi, 2014a; Martin et al., 2022). The second approach involves the use of pooled data from various time periods to capture variations over time. For instance, Agasisti et al. (2014) and Vicente et al. (2021) employed PISA data from 2000 to 2012 and 2003 to 2018, respectively, to explore the impact of education expenditure on academic resilience over time.

Despite the inherent limitation in the sample size of disadvantaged students within the data derived from ILSAs, the statistical power of the aggregated trend analysis can be increased by expanding the number of clusters. However, associations between academic resilience and protective factors were contingent on various covariates and levels considered in the analysis.

Furthermore, the incorporation of data from ILSAs across multiple countries into a pooled dataset has the potential to complicate the interpretation of results due to considerable variations in cultural and educational contexts. Additionally, the inclusion of data from diverse cycles of ILSAs introduces heterogeneity, including educational reforms and policies, which could obscure the findings. As a result, while the approach of aggregated trend analysis offers a broad perspective for exploring academic resilience, it may not comprehensively capture the contextual variations, thereby possibly obscuring or misinterpreting the significance of the findings.

4.2.2 Country-specific analysis

Due to the heterogeneity among nations, some researchers have employed country-specific data to investigate academic resilience across different countries. It is customary for scholars to apply multivariate regression analysis to multiple nations, utilizing the same predictors to enable the examination of the relationship between academic resilience and protective factors across nations. Despite providing education policy-makers an overview of academic resilience across the world, studies using country-specific analysis across numerous countries usually find inconsistent results. To address this issue, certain researchers have chosen to restrict their investigations to countries that share common characteristics, such as culture, language, geographic location, or economic development levels.

Although country-specific analysis takes into account the differences between nations, it presents two primary methodological challenges for researchers. Firstly, the combination of a small cluster size with a limited number of clusters can pose statistical power issues. Secondly, applying the same model across multiple countries may not adequately capture the country-specific relationships between predictors.

Article 2 addressed the limitation of country-specific analysis by examining protective factors at the individual level, where empirical studies have identified more consistent results in the literature. Article 3 utilized latent profile analysis to address the limitations of the two aforementioned approaches. The method incorporated country-specific characteristics into the modeling process, without dividing the data into countries, thus retaining statistical power.

4.3 Testing measurement invariance for protective factors

Measurement invariance ensures consistent interpretation of the construct across groups or time, allowing for meaningful comparisons (Putnick & Bornstein, 2016). If measurement invariance does not hold, the construct is not understood the same way and thus cannot be reasonably compared (Van De Schoot, Lugtig, & Hox, 2012).

In international studies utilizing ILSAs data, academic resilience is commonly operationalized through the utilization of two key criteria: the SES index and student performance. As a result, the need to establish measurement invariance arises not for the construct of academic resilience itself, but rather for its protective factors. However, it is worth noting that measurement invariance is rarely discussed in the field, primarily due to methodological complexities associated with establishing scalar invariance using ILSAs data, as well as researchers' inclination towards composite scale scores as opposed to latent constructs.

4.3.1 Measurement invariance challenges in ILSAs

Measurement invariance is commonly assessed through two methods, namely item response theory (IRT) and confirmatory factor analysis (CFA), with the latter being more prevalent in the context of ILSAs. Within the framework of IRT, measurement invariance is evaluated by examining the constancy of item response functions, which specify probabilities of achieving a score on a test item given a person's latent trait level across groups or time (Kim & Yoon, 2011). Conversely, in CFA, the assessment of measurement invariance involves scrutinizing the equivalence of factor loadings, intercepts/thresholds, and residual variances within a factor model designed to measure an underlying construct (Van De Schoot, Lugtig, & Hox, 2012).

However, previous empirical studies have revealed that achieving scalar invariance, which is necessary for meaningful cross-group comparisons, can be a challenging task due to significant obstacles associated with ILSAs data (Pokropek, Davidov, & Schmidt, 2019; Rutkowski & Svetina, 2014). The challenges primarily encompass three key aspects, specifically the inclusion of extensive sample sizes, a substantial number of nations, and the utilization of categorical response formats (Davidov, Muthen, & Schmidt, 2018).

Considering the substantial number of groups and the sizeable sample size, Rutkowski and Svetina (2014) recommend adopting more relaxed cutoff criteria. For instance, a 0.01 change in the Comparative Fit Index (CFI) and a 0.015 change in the Root Mean Square Error of Approximation (RMSEA) are suggested when performing multigroup confirmatory factor analysis (MG-CFA) on ILSAs data. In order to tackle challenges associated with categorical responses, Svetina and colleagues (2020) introduced the utilization of a threshold model within the framework of MG-CFA. This proposed approach involves conducting equivalence tests for thresholds prior to examining factor loadings, thus offering an optimal methodology specifically tailored for Likert-scaled items.

Furthermore, scholars have also devised novel techniques for assessing measurement invariance, with one of the prominent approaches being the alignment methods proposed by Asparouhov and Muthén (2014c). The alignment method provides a means to estimate group-specific factor means and variances, allowing for the estimation of parameters without requiring exact measurement invariance across multiple groups. The application of this method has demonstrated notable advantages, as evidenced in the existing literature (Munck, Barber, & Torney-Purta, 2018).

4.3.2 Latent variables vs. composite scale scores

In the publicly available data files from ILSAs, a number of composite scale scores are included, in addition to item responses. For example, TIMSS 2019 dataset included principals' ratings of 11 items for the scale of school emphasis on academic success (SEAS). And these 11 items were used to calculate the scale score for each individual through a series of statistical procedures such as partial credit IRT scaling. To demonstrate the cross-country comparability of the context questionnaire scales, TIMSS conducted a principal component analysis of the scale items and presented reliability coefficients for each education system in its technical report (Martin et al., 2020).

The inclusion of multiple predictors may pose significant challenges for scholars in establishing measurement invariance for each latent variable across diverse countries, particularly with conventional methods such as MG-CFA. Consequently, the majority of resilience studies have adopted composite scale scores instead of latent variables in their research, ignoring the issue of measurement invariance (Davidov et al., 2018).

As a result, structural equation modeling (SEM), a statistical approach commonly employed in research treating academic resilience as a personal trait, is seldom employed in studies utilizing ILSAs data. To illustrate, only a few investigations have employed SEM to examine academic resilience in a particular nation, sidestepping the issue of establishing measurement invariance for latent constructs. For instance, Jang et al. (2023) employed SEM to investigate the relationship between academic resilience and students' reading engagement, motivation, and strategies in the United States, based on PISA 2018 data.

Given the number of protective factors and education systems involved in the analyses, empirical studies in this thesis utilized both latent variables and composite scale scores. Specifically, Article 2 explored the influences of two individual protective factors on academic resilience across three economies, with the examination of these two factors being conducted as latent variables. Article 3 investigated 12 protective factors across 36 education systems. To handle the considerable number of factors under investigation, composite scale scores obtained from TIMSS were utilized, with the exception of the latent variable of cognitive activation, for which no corresponding composite scale score was available within the TIMSS dataset.

Empirical studies within this thesis employed Multigroup Confirmatory Factor Analysis (MG-CFA). In Article 2, concerning two latent variables, namely students' sense of belonging to the school and absence from the school, scalar invariance was established by employing a relaxed criterion that allowed for a 0.01 change in Comparative Fit Index (CFI) and a 0.15 change in Root Mean Square Error of Approximation (RMSEA). In Article 3, regarding the latent variable of cognitive activation, scalar invariance was achieved by utilizing the threshold model (Svetina et al., 2020) and employing a relaxed criterion for model fit.

4.4 Multilevel modeling

Multilevel models have become a popular analytical technique for dealing with intricate data structures that display a hierarchical or clustered character. However, the hierarchical structure has not been consistently addressed in resilience research using ILSAs data. Erberer et al. (2015) and Sandoval-Hernández and Bialowolski (2016) investigated protective factors from both individual and school levels, including students' valuing of mathematics and school emphasis on academic achievement. However, they employed single-level logistic regression models that overlooked the hierarchical structure of ILSAs data. Several scholars have recognized the limitations of using single-level regression to study protective factors across different levels, and have consequently turned their attention to investigating protective factors within specific levels. As an illustration, Cheung (2017) conducted research at the individual level, examining the association between academic resilience and variables such as family structure, student self-efficacy, and mathematics anxiety.

4.4.1 Hierarchical considerations

The utilization of ILSAs data has prompted a growing number of scholars to incorporate hierarchical structures in their analytical approaches. This entails considering the nesting structure in one-level regression analysis, which involves accounting for cluster information. A case in point is the work of Özcan and Bulus (2022), who conducted a multi-group logistic regression analysis on two distinct groups of countries classified as having either individualist or collectivist cultural orientations. In the study, they analyzed protective factors at both individual and school levels, such as students' enjoyment of reading and schools' disciplinary climate, by incorporating cluster information in their models to account for the nesting structure. A conventional approach for implementing this technique is to incorporate cluster information into a one-level analysis.

Researchers studying academic resilience using hierarchical data typically use stepwise multilevel models as their primary methodology. This approach entails fitting a series of models beginning with a baseline model containing only student-level variables and progressively incorporating classroom or school-level factors until a final model comprising variables from all levels is attained (Agasisti et al., 2018). This process facilitates the examination of each level's contribution to the outcome and aids in identifying the most critical factors related to academic resilience.

4.4.2 Interactions and multicollinearity

Earlier studies on resilience often concentrated on the interaction between protective factors and demographic variables, such as gender, ethnicity, and immigrant status (Borman & Overman, 2004). With the advancement of MLM, these demographic variables are commonly incorporated into the models as control variables. Consequently, the emphasis on interactions has shifted to exploring the interplay between other variables such as students' psychological factors (i.e., motivation) and contextual factors (i.e., teacher-student relationship).

However, the literature has insufficiently addressed the covariances between protective factors and their interactions across various levels. Scholars have predominantly focused on reporting the relationships between independent variables and the outcome, with little attention to the details of covariance. One plausible explanation for this phenomenon could be attributed to the incorporation of numerous protective factors at different levels, which presents a formidable obstacle in scrutinizing the interplay among covariates. Additionally, if identical multilevel models are implemented across multiple countries with heterogeneous contexts, the inclusion of covariances presents a significant challenge to achieving model convergence across countries. For example, Garcia-Crespo et al. (2021) employed MLM to examine the impact of 10 protective factors at the individual level and 14 protective factors at the classroom level across 23 countries, without delving into any interactions. Likewise, studies that explore multiple protective factors often omit the reporting of correlation coefficients, a crucial process in detecting potential multicollinearity issues, with rare exceptions (Garcia-Crespo et al., 2021).

4.4.3 Protective factors rated by multi-level informants

ILSAs data commonly rely on questionnaires that are administered to a diverse range of stakeholders, including students, parents, teachers, and principals. As a consequence, the data obtained from such assessments comprise information that originates from various sources and levels. Furthermore, the measurement of certain fundamental constructs relies on the input of multiple informants, making it imperative to select an appropriate level for the incorporation of these constructs into multilevel models. For example, Agasisti and Longobardi (2014b) employed PISA data to investigate the relationship between academic resilience and students' absenteeism from principals' perspectives, while Agasisti et al. (2018) investigated the association utilizing students' responses. Embracing diverse perspectives is essential for a more comprehensive understanding of the complex relationship between academic resilience and protective factors. However, this approach also demands careful attention to theoretical and methodological considerations, especially when conducting MLM analyses.

ILSAs data, particularly TIMSS and PIRLS, frequently incorporate two sets of protective factors, namely school climate and instructional quality, which are reported by multiple informants. In TIMSS and PIRLS data, the assessment of school climate involves several dimensions, including disciplinary climate, safe environment, and school emphasis on academic success. Disciplinary climate (i.e., students' absenteeism) is reported by principals, while the safe environment is reported by both teachers and principals, with distinct item sets. However, most of the items related to school emphasis on academic success are similar and are evaluated by both teachers and principals (Martin et al., 2020). Furthermore, as TIMSS and PIRLS pay particular attention to students' curricular and instructional experiences, protective factors associated with instruction are assessed using distinct items from the perspectives of both students and teachers.

Consequently, in multilevel modeling, the appropriate utilization of protective factors at the classroom or school level may involve aggregating ratings provided by students or teachers (Marsh et al., 2012). It should be emphasized that when the number of clusters is below 200, as is often the case in ILSAs data, the Standardized Root Mean Square Residual (SRMR) values at the between level may be high in M*plus* software, despite the model demonstrating a reasonably good fit (Asparouhov & Muthén, 2018).

Within this thesis, Article 3 examined protective factors operating at both the classroom and school levels, involving assessments from students, teachers, and principals. Accordingly, school-level factors, such as the perception of a safe environment by teachers, were aggregated at the school level, while teaching-related factors, such as instructional clarity as reported by students, were aggregated at the classroom level. Following the appropriate treatment of all 12 protective factors, Article 3 employed an SES index to identify disadvantaged students, upon which subsequent analyses were carried out.

4.4.4 Weights application in multilevel analyses

With the increasing application of ILSAs data, more scholars have incorporated sampling weights in their analytical procedures. However, only a small number of them have appropriately addressed this issue in the field of academic resilience (Özcan & Bulus, 2022). ILSAs such as PISA, TIMSS, and PIRLS calculated a series of survey weights to ensure the validity of assessment results and to facilitate fair and precise comparisons of student performance across different countries.

Despite the absence of consensus regarding the implementation of ILSAs weights, it is generally regarded as unsuitable to employ student weights without recalculation, such as the total student weight (*TOTWGT*) in TIMSS, in multilevel modeling (Rutkowski et al., 2010). These weights are more appropriate for single-level analyses (Mullis & Martin, 2017). Rutkowski et al. (2010) recommended the decomposition of the student weights and the application of the appropriate weight at each level. Nonetheless, this approach has received limited attention in the majority of resilience studies utilizing ILSAs (Garcia-Crespo et al., 2021; Jin et al., 2022).

In this thesis, Article 2 performed a one-level analysis while taking cluster information into account. Consequently, the final student weight (*W_FSTUWT*) derived from the PISA data was utilized without recalculation.

4.5 Latent class analaysis

4.5.1 Challenges in linking latent class membership with external variables

LCA is a statistical procedure used to identify latent subpopulations within a sample based on patterns of responses to observed variables (Weller, Bowen, & Faubert, 2020). Specifically, latent class analysis is utilized in instances where the observed variables are categorical. Conversely, when the indicators are continuous, LPA is employed.

Through the examination of response patterns obtained from survey respondents, LCA enables researchers to discern latent subgroups or classes within a given population that demonstrate distinctive attributes pertaining to academic resilience, including students' psychological attributes, family background, and the resources accessible within educational institutions. By means of LCA, scholars can attain a more profound comprehension of the various manifestations of academic resilience profiles present among the population. This valuable information can be utilized to inform the development of targeted interventions and support strategies that are custom-tailored to address the specific requirements of each subgroup.

LCA has experienced a growing utilization within the field of resilience studies in the past decade. Scholars usually employed protective factors, particularly individual psychological characteristics such as self-esteem and efficacy, to identify latent classes of resilient students (Atman-Uslu, 2022). Upon the identification of these latent classes, scholars proceeded to conduct additional investigations on the associations between students' class membership and external variables, including behavior, motivation, and performance (Anthony & Robbins, 2013; Luo et al., 2022). Commonly employed techniques in such inquiries encompass regression analysis or ANOVA. This involves assigning a unique class number to each student and investigating the association between class membership and the external variables under consideration. For example, Luo et al. (2022) utilized ANOVA to investigate the linkage between students' depressive symptoms and their latent class memberships. Similarly, Atman-Uslu (2022) employed ANOVA to establish the correlation between students' latent class membership and their performance and self-efficacy.

However, LCA assigns individuals to classes based on their probability of belonging to classes, given their scores' pattern on the indicator variables (Weller et al., 2020). As a result, the reliability of class allocation is not always ensured. Therefore, the method of assigning a specific class number to individuals tends to overlook the influence of measurement errors and fails to consider the probability of class membership.

4.5.2 The application of the three-step and BCH methods in resilient studies

In order to address concerns associated with measurement errors when connecting latent class membership with external variables, Asparouhov and Muthén (2014a) introduced a three-step method for conducting latent class analysis. This approach involves estimating the latent class measurement model and subsequently examining the association between latent classes and auxiliary variables while accounting for measurement error.

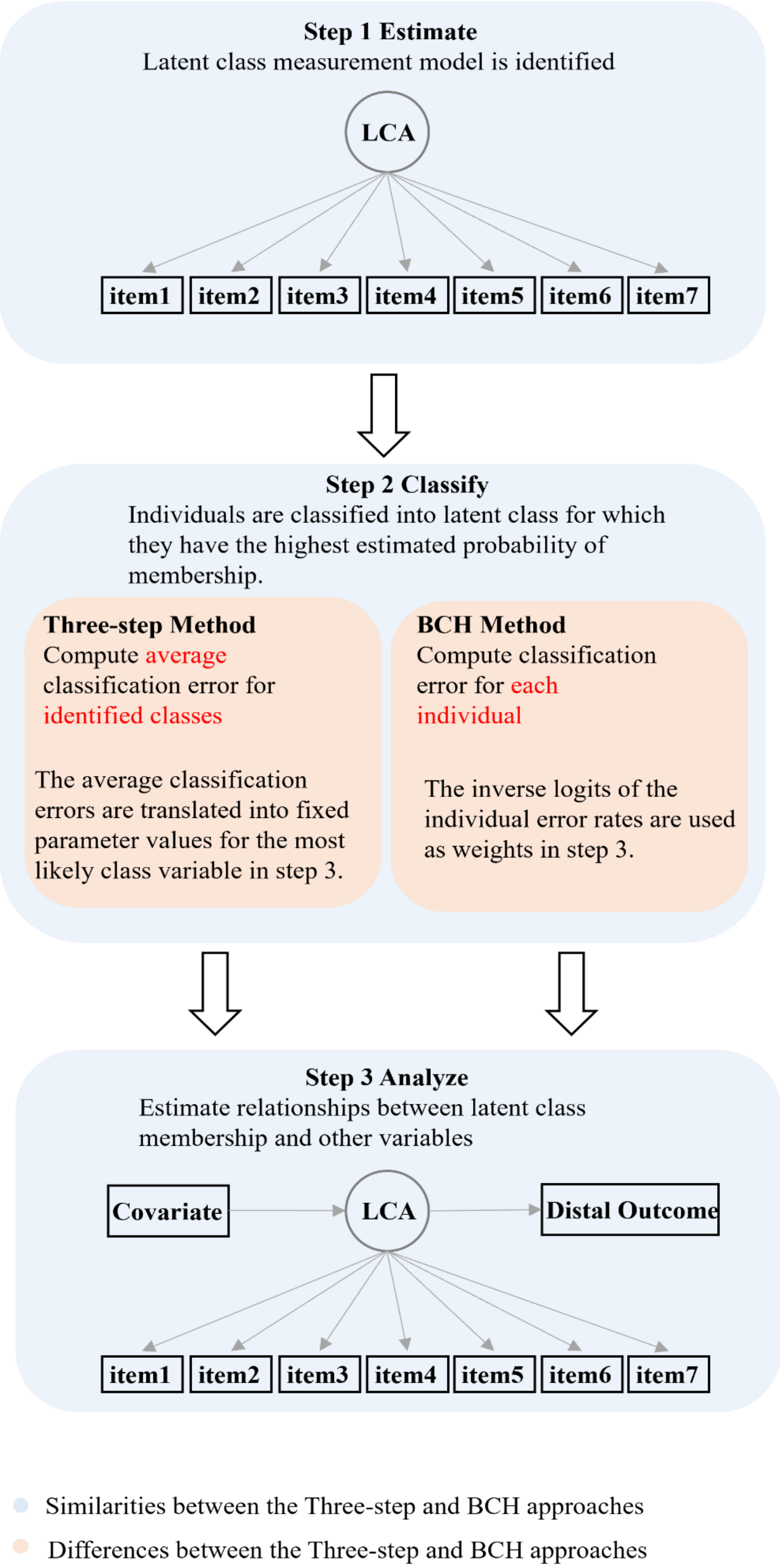
Specifically, in the first step, indicator variables are used to identify the best-fitting latent class model. Once the model is estimated, the posterior probabilities of class membership are calculated, and the modal class assignment is identified. In the second step, the conditional probabilities of a modal class assignment given true latent class membership are computed. In the third and final step, a new analytic model is specified. The most likely class is associated with covariates and distal outcomes, while adjusting for the classification errors obtained in the second step (see Figure 4).

However, the potential for the shift in latent class membership still exists within the three-step method, as the second step computes the average classification error within the sample, while the third step assumes that this error is applied uniformly across all individuals (Nylund-Gibson, Grimm, & Masyn, 2019). In addition, the three-step method may not comprehensively address shifting classes when entropy[3] is low and there is a considerable variance discrepancy in the distal outcome across classes (Bakk, Oberski, & Vermunt, 2014). To address this problem, Asparouhov and Muthén (2014b) proposed the BCH[4] method, which avoids shifts in latent class in the final stage to which the three-step method is susceptible to. The BCH method shares significant similarities with the three-step method. However, it deviates in the second step, where individual classification errors are calculated instead of computing the average classification error. In the third step, the inverse logits of the individual-level error rates are utilized as weights, as opposed to relying on the modal class assignment as an imperfect latent class indicator (Nylund-Gibson et al., 2019). In M*plus* software, the BCH weights are specified by using the "TRAINING" option of the "VARIABLE" command. It is worth noting that the three-step method is compatible with multilevel design in M*plus*, but not the BCH method. However, it is possible to consider cluster information in conjunction with the "TRAINING" option in the BCH method.

[3] Entropy is a statistical fit index for model-based calssification accuracy, with higher values indicating more precise assignment of individuals to latent profiles (Wang et al., 2017). Generally, a value close to 1 is ideal and above .8 is acceptable.

[4] Named after Bolck, Croon and Hagenaars who developed this method.

Figure 4 *Three-step and BCH Methods*



By employing the three-step and BCH methods, researchers are empowered to explore the impact of distinct protective factors, such as school climate, on diverse groups of individuals possessing varying resilient resources, such as individual psychological characteristics. As a result, there has been a growing adoption of the three-step approach in academic resilience investigations, as evidenced by studies conducted by Boutin-Martinez et al. (2019), Lines et al. (2020), and Koirikivi et al. (2021).

Although the three-step method has gained popularity in research utilizing ILSAs data (Wu et al., 2021), to the best of my knowledge, there has been no inquiry that applies this method to investigate academic resilience in the context of ILSAs data. Besides, for research that adopted the three-step method, only a limited number of studies have addressed the issue of hierarchical structure (Mäkikangas et al., 2018; Teig & Nilsen, 2022). Given the challenges associated with guaranteeing invariance of both the measurement model and the latent class distribution across classes, further theoretical and methodological research is needed to facilitate exploration in a hierarchical context.

In order to address the limitations associated with MLM stepwise approach in international studies, Article 3 employed the latent profile analysis (LPA) with a three-step BCH method to investigate profiles of resilient resources. The study subsequently examined the presence of these profiles across six cultural groups and explored the relationship between education expenditure and academic resilience across these profiles.

5 Summary of the articles

The main objective of this doctoral thesis was two-fold. Firstly, it aimed to explore how academic resilience can be operationalized to meaningfully identify resilient students and analyze factors protecting students from their adversities across different countries. Secondly, it aimed to examine the methods that can be utilized to investigate academic resilience at a global level. These two aims have been comprehensively examined in three distinct articles co-authored with other researchers. The present chapter provides a summary of each article. Figure 5 illustrates the interconnections between the three articles. In particular, Article 1 examined various issues encompassing protective factors, the operationalization of academic resilience, data considerations, and research methodologies. These four perspectives are graphically depicted in Figure 5 using the colors magenta, blue, green, magenta, and orange, respectively. Subsequently, Articles 2 and 3 further delved into and elaborated on the identified issues.

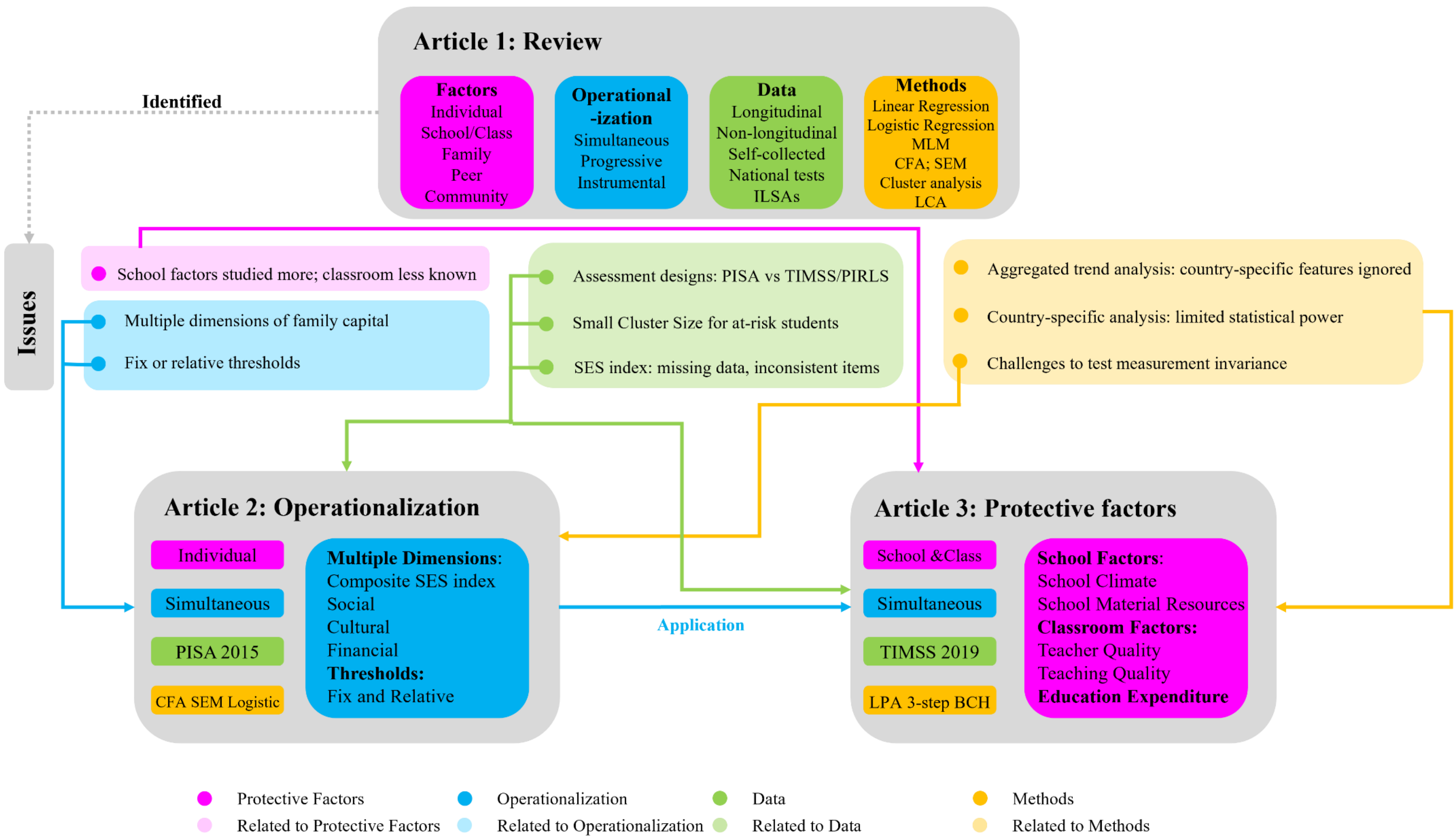
5.1 Article 1: Review

Ye, W., Teig, N., & Blömeke, S. (2023). Systematic review of protective factors related to academic resilience in children and adolescents: Unpacking the interplay of operationalization, data, and research method (under review in the journal of Educational Research Review)

Article 1 conducted a systematic review of protective factors associated with academic resilience among school-aged students. Typically, studies on resilience in psychology and sociology concentrate on children and adolescents, and the inquiry into protective factors that promote resilience is frequently grounded in these fields. However, the exploration of the domain of education in this context has been relatively less examined.

The exploration of protective factors that promote academic resilience has been an area of significant interest in the field. However, the heterogeneity observed in the operationalization of academic resilience, the sources of data used, the timeframes considered, and the research methods employed have impeded the understanding of the precise extent to which the protective factors identified in prior studies truly contribute to the outcomes observed, or if they are confounded by these variances. In order to clarify these inquiries, Article 1 analyzed five distinct groups of *protective factors* (individual, family, school, peer, and community), in conjunction with three types of *operationalizations* for academic resilience (simultaneous, progressive, and instrumental), two timeframes (longitudinal and

Figure 5 *Relationships Among Three Articles*



non-longitudinal), three types of *data sources* (self-collected, national/local assessments, ILSAs), and commonly employed *research methods* in 119 empirical studies.

Factors

The two most extensively researched categories of protective factors among the five identified groups are those relating to individual and school-level characteristics. The former is often subjected to "instrumental" operationalization and structural equation modeling, while the latter is typically examined through "simultaneous" or "progressive" operationalizations and multilevel modeling techniques. In the literature, the impact of examined protective factors has yielded both consistent and inconsistent findings. These divergent results can be attributed, in part, to the fact that different measurement instruments and different operationalizations of resilience are used across studies. Moreover, discrepancies across studies partly reflect that several studies do not have sufficient statistical power to establish statistical significance.

Operationalizations

The study identified three operationalizations of academic resilience, namely simultaneous, progressive, and instrumental. The first two operationalizations are consistent with the theoretical perspective that treats academic resilience as a dynamic interplay between individuals and their contexts. On the other hand, the third operationalization regards academic resilience as an inherent personal trait. As a result, diverse interpretations of academic resilience have given rise to variations in the selection and application of protective factors, data, and research methods.

Data

The examination of 119 studies revealed that a significant proportion of them, namely, approximately 31% and 16%, utilized national assessments and ILSAs data, respectively. These data sources were found to promote the exploration of school-level factors in the field, with the former facilitating the exploration of protective factors across time and the latter contributing to the investigation of teaching-related factors.

Research Methods

Academic resilience can be operationalized as a continuous, binary, or latent variable, with corresponding statistical techniques being linear regression, logistic regression, and structural equation modeling. As the research into school-level protective factors continues to expand, MLM has become increasingly prevalent. Furthermore, researchers have utilized cluster analysis and latent class analysis in this field.

In sum, the present review contributes to the extant literature by addressing a research gap in the examination of protective factors in education. Moreover, the study establishes a connection between protective factors, operationalization, data, and research methods. Additionally, this review identified and discussed the challenges and concerns relating to these four perspectives in empirical studies.

5.2 Article 2: Operationalization

Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. *Educational Assessment, Evaluation and Accountability, 33*(1), 169-202. https://doi.org/10.1007/s11092-020-09351-7

Article 2 focused on the operationalization of academic resilience, specifically narrowing its focus to the "simultaneous" approach utilized in international studies that employ ILSAs data. The historical underpinnings and subsequent evolution of academic resilience have traditionally limited its examination to a homogenous context, confined within a particular country. Consequently, international investigations into academic resilience encounter diverse challenges related to the reliability and validity of the construct.

In international studies, especially those utilizing ILSAs data, the concept of academic resilience is commonly assessed by two primary components: the socio-economic status (SES) of students and their academic performance. With the growing recognition of domain-specific variations in academic resilience, some scholars have employed performance across mathematics, science, and reading to identify high-achieving students (Gabrielli et al., 2022). However, there has been limited attention given to identifying disadvantaged students. Most studies employing ILSAs data have employed a combined SES index to identify disadvantaged students, yet the specific components of this composite index remain largely unexplored. Furthermore, the thresholds for the two components, namely students' SES backgrounds and performance, have varied across studies, making cross-study comparisons challenging.

This study utilized PISA 2015 data from three diverse economies, namely Norway, Hong Kong, and Peru, to address the following research questions:

1. How large is the group of academically resilient students when different conceptualizations of academic resilience are applied?
2. How do these conceptualizations of academic resilience affect which students are classified as academically resilient when it comes to gender and language background?

3. How are different conceptualizations of academic resilience associated with external variables, which can be supposed to assess similar constructs?
4. Do results change if different indicators of students’ capital (economic, social, and cultural) are used?

In order to identify students who are at risk, four background indicators were employed, namely, a composite SES index and economic, social, and cultural capitals. Additionally, two thresholds, specifically fixed and relative thresholds, were combined with the students’ mathematics performance, resulting in 16 distinct operationalizations. These operationalizations were examined in conjunction with two demographic variables, namely gender and language, as well as two individual protective factors, namely students’ sense of belonging and absence from school.

This study found that when a fixed background threshold was applied, the classification was likely to be affected by the developmental state of a country. Similarly, the classification result was substantially influenced by the performance level of the country when a fixed performance threshold was implemented. As such, the adoption of fixed thresholds may result in over- or under-estimating academically resilient students in certain countries.

Moreover, the composition of academically resilient students varied significantly by gender and language depending on which indicator of human capital or which thresholds were applied, reflecting underlying societal characteristics. Conclusions drawn from varying results based on diverse conceptualizations and operationalizations would exhibit significant variations. Additionally, compared to the application of a social or economic capital indicator, applying a cultural capital indicator may lead to lower shares of disadvantaged students classified as academically resilient.

Furthermore, the associations between academic resilience and the two distal factors (students’ sense of belonging and absenteeism) exhibited contextual variation. A stronger sense of belonging to school significantly increased the chances of being classified as academically resilient in Peru, but not in Norway or Hong Kong. Conversely, absenteeism was linked to resilience in Norway and Hong Kong, but not in Peru.

In sum, the present article undertook a comparative analysis of different operationalizations and addressed concerns about the thresholds employed in international research. By investigating four background indicators, this study has contributed to a better comprehension of the impact of multiple capital dimensions on academic resilience in the literature.

5.3 Article 3: Protective Factors

Ye, W., Olsen, R. V., & Blömeke, S. (2023). More money does not necessarily help: Relations of education expenditure, school characteristics, and academic resilience across 36 education systems (under review in the journal of Large-scale Assessments in Education)

Article 3 aimed to investigate the relationship between academic resilience and protective factors such as teacher quality, teaching quality, school resources, school climate, and education expenditure. Building upon the operationalization of academic resilience developed in Article 1, this study aimed to address the under-researched area of classroom-level factors in the existing literature. Acknowledging the significant influence of classroom-level factors on students' learning experiences and academic performance, the research utilized data from TIMSS 2019, which encompasses 36 education systems. This study specifically emphasized teachers and their instructional practices, extending the analysis beyond school-level protective factors.

Given the inherently limited cluster size of disadvantaged students in ILSAs data, this study adopted latent profile analysis (LPA) to identify profiles of school and classroom protective factors that contribute to academic resilience. During the process of identifying the most appropriate latent profile model, this study took into account the covariances among protective factors, which allowed for the consideration of unique characteristics specific to each education system. Additionally, this method entails analyzing information across the 36 educational systems collectively, rather than individually, thereby circumventing substantial statistical power loss.

The present study aimed to address the following inquiries through the utilization of LPA in conjunction with a three-step BCH approach:

1. How many distinct profiles of resilience resources, characterized by teacher quality, teaching quality, school climate, and school resources, can be identified in the sample?
2. Do the profiles of resilience resources exhibit identifiable cultural patterns across diverse nations?
3. To what extent do the identified latent profiles predict academic resilience?
4. To what extent do the associations between education expenditure as a percentage of Gross Domestic Product (GDP) and academic resilience vary across the identified profiles?

This study identified four resilient resource profiles based on 11 protective factors related to teacher quality, teaching quality, school resources, and school climate. The first profile, named "*Vulnerable*," had the lowest levels of teacher quality, teaching quality, school resources, and school climate. The second profile, named "*Effective Teaching and Positive Climate*," had high levels of teaching quality and school climate. The third profile, named "*Resource-Heavy, Quality-Light*," had high levels of school resources but low levels of teacher and teaching quality. The fourth profile, named "*Good Schools*," had high levels of teacher quality, teaching quality, school resources, and school climate.

Further, this study explored the presence of these four profiles within six cultural groups, namely, Confucian Asia, Middle East, Post-Soviet, Latin Europe, Anglo, and Nordic countries. The majority of the cultural groups examined exhibit a degree of cultural similarities, as evidenced by no significant differences in the respective profiles of protective factors in five out of twelve Middle Eastern countries, four out of five Confucian Asian economies, three out of five Anglo countries, two out of three Nordic countries, and two out of six Post-Soviet countries. However, differences are evident within these cultural groups, particularly in the Middle East and Post-Soviet countries. This underscores the importance of contextual considerations in international studies, as differences appear to be explained by variations in economic development.

Moreover, this study explored the predictive capacity of these four profiles concerning academic resilience. Results indicated that students in Profile 4 ("*Good Schools*") exhibited significantly higher resilience than those in the other three profiles.

Additionally, this study examined the association between academic resilience and education expenditure across these four identified profiles. Results revealed a non-significant negative association between education expenditure and academic resilience in Profile 1. In contrast, positive associations were observed in the other three profiles, with statistically significant relationships found only in Profiles 2 ("*Effective Teaching and Positive Climate*") and 3 ("*Resource-Heavy, Quality-Light*").

In sum, this paper explored new approaches for assessing academic resilience using ILSAs data in international studies. Furthermore, the association between educational expenditure and academic resilience has exhibited significant variations contingent upon the profiles of protective factors in schools and classrooms, thereby underscoring the relevance of contextual influences in exploring academic resilience.

6 Discussion and implication

This doctoral research aimed to explore the operationalization of academic resilience in international studies and methods for investigating protective factors using ILSAs data. This chapter summarizes how this thesis, including the three articles, has contributed to these two aims in terms of theoretical, empirical, and methodological advancements. Next, this chapter examines the implications of these findings and highlights the issues that require attention for researchers, ILSAs, and policymakers. Finally, this chapter discusses the strengths and limitations of using ILSAs data to investigate academic resilience, followed by a brief concluding remark.

6.1 Theoretical, empirical, and methodological contributions

6.1.1 Theoretical contributions

The systematic review conducted in Article 1 plays a crucial role in contributing to the theoretical framework of academic resilience. By synthesizing existing literature and examining multiple perspectives, this review contributes to a comprehensive understanding of four key areas: (1) the investigation of protective factors that facilitate the development of academic resilience, (2) the measurement approaches employed to assess academic resilience, (3) the utilization of data sources in studying academic resilience, and (4) the research methodologies adopted to explore the intricate relationship between protective factors and academic resilience. By encompassing these four perspectives, this review offers an in-depth analysis and a holistic view of the multifaceted nature of academic resilience research.

In Article 2, a thorough examination of various dimensions of family capital was conducted. The findings revealed that utilizing cultural capital as a means to identify disadvantaged students might result in a decreased representation of academic resilience. Additionally, the study identified that the utilization of fixed thresholds could lead to an inaccurate estimation of academic resilience, either underestimating or overestimating its prevalence across different countries. Through the comparison of coefficients between external variables and academic resilience across 16 operationalizations, the study revealed minimal significant differences. These findings collectively contribute to the theoretical comprehension surrounding the measurement of academic resilience in the context of international studies.

6.1.2 Empirical contributions

The empirical studies included in this thesis, namely Articles 2 and 3, built on the systematic review in Article 1 to examine academic resilience within the context of international studies, employing data from ILSAs. Article 2 focused on issues related to the measurement of academic resilience, while Article 3 delved into the exploration of the impact of protective factors.

The most significant challenge encountered when investigating academic resilience within international studies revolves around determining the thresholds for risks and positive adaptations. Article 2 employed fixed and relative thresholds to analyze their impact on academic resilience in international studies. Fixed thresholds involve the application of the same cut-off score across educational systems, while relative thresholds involve the application of different cut-off scores. The findings indicated that the application of a fixed threshold in identifying academically resilient students is influenced by a country’s economic development level and academic performance level. As exemplified in Article 2, the implementation of a fixed performance threshold results in the identification of 91.85% of disadvantaged students who demonstrate academic resilience in the context of Hong Kong.

Moreover, Article 2 contributed to the existing literature by examining multiple dimensions of family background, including a composite SES index, and economic, social, and cultural perspectives. The findings indicated that the composition of resilient students varies when employing different background indicators as indicators of adversity, emphasizing the importance of the nature and composition of the SES index in academic resilience research. Specifically, the study revealed that the inclusion of cultural background indicators may lead to a decrease in the proportion of students classified as academically resilient.

Article 3 mainly addressed issues relating to the influence of protective factors on academic resilience. To underscore the significance of the classroom environment, Article 3 examined protective factors pertaining to school climate, school resources, teacher quality, and teaching quality. Given the exploration of academic resilience in countries with comparable cultural backgrounds by several scholars, Article 3 conducted a further investigation into the commonalities and distinctions among six cultural groups. This examination resulted in the identification of both similarities and differences.

Based on previous findings on the influence of education expenditures on academic resilience (Agasisti et al, 2018), Article 3 further explored this relationship through the profiles of resilience resources. Consistent with prior research, the results demonstrated that students attending schools characterized by relatively high levels of school and classroom protective factors are more likely to exhibit resilience. In schools with comparatively lower levels of teacher quality and school climate, education expenditure demonstrated a significant association with academic resilience. Moreover, the analyses suggest that for schools with the lowest levels of school climate, school resources, teacher quality, and teaching quality, education expenditure failed to predict academic resilience.

6.1.3 Methodological contributions

Article 3 focused on the methodological challenges associated with the two most prevalent approaches used in international studies, namely, the aggregated trend analysis and country-specific analysis. The former approach overlooks country-specific information, whereas the latter may fail to yield significant findings due to inadequate statistical power. Similar to previous studies, Article 3 examined numerous protective factors and encompassed a wide range of countries with diverse characteristics. However, Article 3 differed from previous studies by identifying latent profiles with 11 protective factors related to schools and classrooms across 36 education systems. In the process of model identification, the covariances among these 11 indicators were taken into account, which considered the unique characteristics of each education system. After identifying the number of latent profiles, they were associated with a covariate (education expenditure) and the distal outcome (academic resilience) using a three-step BCH method. This approach was employed to avoid the loss of statistical power and did not require the division of the data into country-specific datasets.

6.2 Implications for researchers, ILSAs, and policymakers

6.2.1 Implications for researchers

Simultaneous and progressive operationalizations

The majority of studies utilizing ILSAs data have employed simultaneous operationalization to define academic resilience. Although assessed at the same time point, students' SES backgrounds are typically antecedent to their performance on ILSAs. However, under circumstances where other factors, such as academic setbacks like absenteeism, are regarded as risks, a progressive operationalization may offer greater efficacy. Nonetheless, the employment of a cross-sectional design poses a formidable challenge in disentangling risk and outcome at distinct time points.

One potential strategy to tackle this issue involves utilizing national assessments that are associated with ILSAs. As an example, Thiessen (2008) specified that students at risk are those who demonstrated low reading performance in PISA 2000. Thiessen (2008) then used data from the Canadian longitudinal Youth in Transition Survey, which is an extension of PISA 2000, to assess positive adaptations based on the students' academic achievements four years later. This approach allows for the definition of risk at an early stage, thereby facilitating the exploration of how protective factors influence outcomes over time.

Another alternative approach to operating academic resilience as a progressive procedure is to adopt a national perspective, as demonstrated by Agasisti and Longobardi (2014a), which involves examining the prevalence of academically resilient students across countries or periods. For ILSAs that incorporate a time-lagged cross-sectional design design, such as TIMSS, it is feasible to investigate protective factors related to the presence of academic resilience, both in fourth-grade students and in those who have progressed to eighth-grade, over a four-year period.

Risks from multiple perspectives

Article 2 found that when cultural capital indicators are used to identify disadvantaged students, the likelihood of finding resilient students is lower compared to other indicators. Therefore, further investigation is necessary to better understand how cultural capital affects the probability of being resilient.

Given that some ILSAs may lack information on students' family backgrounds, a possible strategy to uncover the multiple dimensions of family-related risks is to utilize local datasets from government agencies. This would necessitate linking the ILSAs data with the local database. For instance, Fantuzzo et al. (2012) demonstrated the feasibility of this strategy by linking students' academic achievement data obtained from national assessments to the risk factors of child maltreatment and homelessness, which were sourced from the Department of Human Services and the Office of Supportive Housing in the United States, respectively.

An alternative strategy involves targeting the ILSAs participating countries that exhibit comparatively lower levels of missing data related to students' family backgrounds. Specifically, this approach may prioritize economies from East Asia, as they exhibit relatively high response rates for both students and their families (Lam & Zhou, 2021).

Moreover, an investigation of risks encountered in educational settings can enhance a holistic comprehension of academic resilience. Nevertheless, it is noteworthy that the effectiveness of protective factors in mitigating the impact of a risk factor, such as school absenteeism, may require a considerable amount of time. Consequently, utilizing a simultaneous operationalization approach, which evaluates both the exposure to risks and positive adaptation simultaneously, may not be optimal for this research design.

Application of plausible values

Despite the inconsistent application of plausible values in resilience studies, a majority of scholars have acknowledged the significance of utilizing them appropriately. Resilience studies typically involve the conversion of plausible values into binary format and the subsequent utilization of such transformed binary data for analysis. Articles 2 and 3 incorporated all plausible values in their respective analyses, and the coefficients derived from all plausible values and each individual value were frequently comparable.

However, it is noteworthy that the impact of converting continuous data into binary format on the outcome of the analysis is not yet fully comprehended. Therefore, future investigations are recommended to delve into these differences in more detail.

Context considerations for performance

The majority of resilience studies have adopted a context-specific approach in identifying students who are at risk across various nations. However, in the case of defining high-achieving students, country-specific characteristics have received comparatively less attention. One potential explanation for this phenomenon is the provision of a performance benchmark, denoting a specific level of proficiency for students, by many ILSAs such as PISA and TIMSS. For example, several scholars in the field have utilized PISA level 3 (representing moderate proficiency) and TIMSS score 475 (an intermediate international benchmark) to identify resilient students (Erberer et al., 2015; OECD, 2011). The levels of proficiency, however, are derived from the aggregate performance of all participants, and thus may not provide a precise indication of moderate proficiency within a specific country.

More specifically, the academic performance of disadvantaged students is crucial in determining their future success, including but not limited to, attending college, participating in job marketing, and breaking the cycle of poverty or adversity. It is worth noting that, even if a student's academic performance is comparatively lower than that of students in high-achieving countries, the student is more likely to succeed if her performance is better than her peers within the country. Therefore, it is crucial to consider the local context when setting academic standards to ensure a fair evaluation of the academic achievements of disadvantaged students within their country.

While current studies typically use a top 1/3 or 1/4 threshold in performance to identify high-achievers, future studies could consider country-specific information, such as the percentage of students attending university or participating in the job market, to determine appropriate cut-off values. Additionally, researchers could take into account the coverage rate of the target-aged population by the assessment as another relevant factor.

MLM stepwise approach

The MLM stepwise approach, although commonly used in international resilience studies, is constrained by two inherent issues: the inability to account for country-specific interactions and the small cluster size of disadvantaged students. The application of the MLM stepwise approach to pooled data from various countries may not account for country-specific considerations. Conversely, the application of multilevel models to each country is beset by reduced statistical power and model-fitting difficulties such as convergence.

One potential remedy for the issue of limited cluster size is to augment the number of clusters, consequently bolstering the statistical power. A possible approach to achieve this is to pool data from multiple ILSAs cycles for each country, such as utilizing PISA data from 2015 and 2018. Nevertheless, it should be noted that this strategy has a drawback. Specifically, the accumulation of heterogeneous information during different ILSAs cycles, including shifts in education reform and policies, may obscure the associations between protective factors and academic resilience.

In future research investigating academic resilience using the MLM stepwise approach, it is recommended to apply it in a homogenous context with adequate statistical power, such as a large number of reasonably sized clusters. When using this approach to investigate academic resilience across nations, it is advisable to expand the scope of discussion beyond significant results and consider the direction of the relationships between protective factors and academic resilience.

A potential application of the MLM stepwise approach is to investigate the interactions between individuals and their context, such as students' motivation and teachers' expectations. Despite the growing attention to malleable factors associated with schools and classrooms, limited studies have examined the interactions between individuals and their context. The MLM stepwise approach can provide a promising solution for addressing such research questions.

LCA three-Step approach

The latent class analysis represents a viable alternative methodology for investigating protective factors in international research. Through this approach, it becomes feasible to identify latent classes of resources that promote resilience, including individual psychological factors, school resources and climate, as well as teacher and teaching-related factors. By utilizing a three-step approach, these identified classes can be analyzed in relation to external variables, such as demographic factors (i.e., gender), contextual factors (i.e. school location), and academic outcomes (i.e., student achievement).

Nevertheless, the LCA three-step approach is associated with two primary issues. Firstly, although it is possible to establish the connection between external variables and identified latent classes, it is not possible to scrutinize the relationship between external variables and latent class indicator variables. Secondly, the application of the LCA three-step approach in a multilevel context necessitates further development in statistical modeling and empirical studies. The three-step method employs average classification error rather than individual classification error, which may result in potential shifting in class membership during the final stage. The BCH method, as an extension of the three-step approach, addresses the classification error issues related to class membership changes. Both the three-step and BCH methods were utilized in the preliminary analysis for Article 3. However, considering that Profile 1 has a small size and differs significantly from the other three profiles in terms of its relationship with external variables, the BCH method produced more consistent results across five plausible values than the three-step method. Nonetheless, the BCH approach in the M*plus* setting is not compatible with multilevel design, despite accommodating cluster considerations. As a result, with currently available techniques, it is possible to consider no more than two levels of hierarchical structures in the BCH method within the M*plus* setting.

6.2.2 Implications for ILSAs

Academic resilience is typically not directly assessed in studies utilizing ILSA data. Instead, it is inferred through the combination of risk factors and positive adaptations. Protective factors, such as student motivation, parental support, and school climate, are often considered elements that aid disadvantaged students in achieving better academic performance. However, this approach fails to fully acknowledge that some students may possess stronger inherent characteristics than others, enabling them to bounce back more effectively in the face of difficulties. Consequently, the current research approach in ILSAs overlooks a crucial piece of information: the interplay of various personal characteristics as an internal source of resilience in academic settings. Moreover, this internal resilience can be nurtured by external factors, including parental care, teacher support, and friendships. The recognition and identification of this inner power not only contribute to a more comprehensive understanding of the underlying mechanisms between protective factors but also serve as a foundation for cultivating academic resilience.

Academic resilience encompasses two aspects: firstly, it denotes the individual student's ability to achieve academic success despite adversity; secondly, it pertains to the contextual capacity to facilitate students' resilience and recovery. Considering these dual dimensions, adopting multiple perspectives on academic resilience can yield a richer and more comprehensive understanding of the field.

Self-report surveys allow students to articulate their own experiences, perspectives, and opinions. These first-person accounts offer valuable insights into the individual circumstances of students, encompassing both the challenges they encounter (risk factors) and the resources and support systems that aid in their coping mechanisms (protective factors). Given that students themselves directly experience resilience (or the absence thereof), their viewpoints hold immense significance in comprehending this intricate construct. Hence, it is recommended to incorporate items and measures specifically designed to assess students' academic resilience within the framework of ILSAs. Moreover, the incorporation of resilience items within ILSAs can effectively tackle prevailing issues, such as the limited cluster size of disadvantaged students and the inherent challenges in conducting international comparisons across diverse countries.

6.2.3 Implications for policymakers

This thesis presents a comprehensive exploration of critical factors that contribute to academic resilience among students, with a specific focus on modifiable elements within educational settings that can be influenced through educational policies. However, empirical studies conducted in this thesis indicate that a uniform approach may not effectively enhance academic resilience in all countries. For instance, findings from Article 2 reveal that the associations between academic resilience and students' sense of belonging and school attendance vary across different countries. Likewise, Article 3 demonstrates that simply increasing education expenditure in schools with ample resources may not lead to significant improvements in academic resilience.

Hence, it is imperative for policymakers to acknowledge that promoting academic resilience necessitates a comprehensive and nuanced approach, considering the specific contextual factors at play. By implementing evidence-based policies and targeted interventions tailored to the distinct needs of each educational setting, we can foster a resilient learning environment that empowers students to overcome challenges and unlock their full potential. It is crucial to recognize that the effectiveness of these interventions may vary across countries and educational contexts. Therefore, policymakers should employ a contextualized approach that takes into account the unique characteristics and challenges of each educational system. By doing so, we can enhance academic resilience and create a supportive and conducive educational environment that fosters optimal student development and success.

6.3 Strengths and limitations of using ILSAs data to investigate academic resilience

The utilization of ILSAs data provides significant advantages in the exploration of academic resilience. ILSAs encompass a wide range of countries and student populations, thus yielding a comprehensive and diverse dataset that includes contextual information. The standardized nature of these assessments facilitates rigorous comparisons across countries, regions, and demographic groups. By conducting regular assessments at defined intervals, such as PISA and TIMSS, it becomes feasible to track trends over time and examine the dynamics of resilience and changes in educational outcomes. Additionally, ILSAs encompass multifaceted evaluations of student knowledge, skills, dispositions, and learning activities. This comprehensive data collection enables the identification of resilience patterns among disadvantaged students within specific countries or regions, thereby potentially uncovering effective practices that can be replicated in or modified to fit other settings. Consequently, the utilization of ILSAs data offers a robust framework for comprehending and fostering academic resilience on a global scale.

However, the use of ILSAs data for investigating academic resilience is not without its limitations. One key drawback is that not all students or schools participate in these assessments, leading to potential sample biases. As a result, the samples may not fully represent the diverse circumstances and backgrounds of all disadvantaged students within a country. While these assessments provide valuable insights into academic resilience, they may not capture other important dimensions of resilience, such as emotional or social resilience, which are integral to a holistic understanding of student well-being. Additionally, ILSAs primarily offer averages and general trends, which may not adequately capture the nuanced systemic inequities present within individual nations' educational systems. For instance, a country with a high proportion of resilient students may suggest a favorable overall resilience rate, but this might obscure significant disparities between different regions or between rural and urban schools. It is crucial to recognize and consider these limitations when utilizing ILSAs data for investigating academic resilience, as they can impact the comprehensive understanding and targeted support for resilient students.

6.4 Concluding remarks

Since the 2010s, there has been a significant increase in the use of ILSAs data in resilience studies, specifically in international research. However, the development of academic resilience from a homogeneous to a heterogeneous context has raised two issues related to operationalization and methods. The primary objective of this thesis was to address these issues by developing appropriate operationalizations of academic resilience for international studies and statistical methods for exploring protective factors.

Regarding the operationalization of academic resilience, this thesis recommended taking into account contextual features when defining risk and positive adaptations. In addition, this thesis addressed methodological challenges, including the decreased statistical power resulting from small cluster sizes, the consideration of level-specific responses from multiple informants, and the testing of measurement invariance. Appropriate applications of weights and plausible values were also discussed. Furthermore, this thesis introduced new methods to explore protective factors and academic resilience, such as the use of LPA with a three-step BCH method.

In summary, this thesis emphasized the significance of accounting for context-specific features in investigating academic resilience in international studies, including both operationalization and methodology.

References

Aburn, G., Gott, M., & Hoare, K. (2016). What is resilience? An integrative review of the empirical literature. *Journal of Advanced Nursing*, *72*(5), 980-1000. https://doi.org/10.1111/jan.12888

Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). Academic resilience: What schools and countries do to help advantaged students succeed in PISA. *OECD Education Working Papers*, *167*. https://doi.org/10.1787/19939019

Agasisti, T., & Longobardi, S. (2014a). Educational institutions, resources, and students' resiliency: an empirical study about OECD countries. *Economics Bulletin*, *34*(2), 1055-1067. Retrieved from https://EconPapers.repec.org/RePEc:ebl:ecbull:eb-14-00160

Agasisti, T., & Longobardi, S. (2014b). Inequality in education: Can Italian disadvantaged students close the gap. *Journal of Behavioral and Experimental Economics*, *52*, 8-20. https://doi.org/10.1016/j.socec.2014.05.002

Agasisti, T., Longobardi, S., & Regoli, A. (2014). Does public spending improve educational resilience? A longitudinal analysis of OECD-PISA data. *Working Papers 3, Società Italiana di Economia Pubblica*.

Anthony, E. K., & Robbins, D. E. (2013). A latent class analysis of resilient development among early adolescents living in public housing. *Children and Youth Services Review*, *35*(1), 82-90. https://doi.org/10.1016/j.childyouth.2012.10.012

Asparouhov, T., & Muthén, B. (2014a). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. (2014b). Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes*, *21*(2), 1-22.

Asparouhov, T., & Muthén, B. (2014c). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495-508. https://doi.org/10.1080/10705511.2014.919210

Asparouhov, T., & Muthén, B. (2018). SRMR in Mplus. Retrieved from http://www.statmodel.com/download/SRMR2.pdf

Atman-Uslu, N. (2022). How do computational thinking self-efficacy and performance differ according to secondary school students' profiles? The role of computational identity, academic resilience, and gender. *Education and Information Technologies*, 1-25. https://doi.org/10.1007/s10639-022-11425-6

Austin, J. L., Jeffries, E. F., Winston III, W., & Brady, S. S. (2022). Race-related stressors and resources for resilience: Associations with emotional health, conduct problems, and academic investment among African American early adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 61*(4), 544-553. https://doi.org/10.1016/j.jaac.2021.05.020

Avvisati, F. (2020). The measure of socio-economic status in PISA: A review and some suggested improvements. *Large-Scale Assessments in Education, 8*(1), 1-37. https://doi.org/10.1186/s40536-020-00086-x

Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2014). Relating latent class assignments to external variables: Standard errors for correct inference. *Political Analysis, 22*(4). https://doi.org/10.1093/pan/mpu003

Bellis, M. A., Hughes, K., Ford, K., Hardcastle, K. A., Sharp, C. A., Wood, S., . . . Davies, A. (2018). Adverse childhood experiences and sources of childhood resilience: A retrospective study of their combined relationships with child health and educational attendance. *BMC Public Health, 18*(1), 1-12. https://doi.org/10.1186/s12889-018-5699-8

Bester, G., & Kuyper, N. (2020). The influence of additional educational support on poverty-stricken adolescents' resilience and academic performance. *Africa Education Review, 17*(3), 158-174. https://doi.org/10.1080/18146627.2019.1689149

Bhamra, R., Dani, S., & Burnard, K. (2011). Resilience: the concept, a literature review and future directions. *International Journal of Production Research, 49*(18), 5375-5393. https://doi.org/10.1080/00207543.2011.563826

Borman, G. D., & Overman, L. T. (2004). Academic resilience in mathematics among poor and minority students. *The Elementary School Journal, 104*(3), 177-195.

Bostwick, K. C. P., Martin, A. J., Collie, R. J., Burns, E. C., Hare, N., Cox, S., . . . McCarthy, I. (2022). Academic buoyancy in high school: A cross-lagged multilevel modeling approach exploring reciprocal effects with perceived school support, motivation, and engagement. *Journal of Educational Psychology, 114*(8), 1931-199. https://dx.doi.org/10.1037/edu0000753

Boutin-Martinez, A., Mireles-Rios, R., Nylund-Gibson, K., & Simon, O. (2019). Exploring resilience in Latina/o academic outcomes: A latent class approach. *Journal of*

Education for Students Placed at Risk, 24(2), 174-191. http://dx.doi.org/10.1080/10824669.2019.1594817

Buckner, J. C., Mezzacappa, E., & Beardslee, W. R. (2003). Characteristics of resilient youths living in poverty: The role of self-regulatory processes. *Development and Psychopathology, 15*(1), 139-162. https://doi.org/10.1017/S0954579403000087

Bussemakers, C., & Kraaykamp, G. (2020). Youth adversity, parental resources and educational attainment: Contrasting a resilience and a reproduction perspective. *Research in Social Stratification and Mobility, 67*, 100505. https://doi.org/10.1016/j.rssm.2020.100505

Cappella, E., & Weinstein, R. S. (2001). Turning around reading achievement: Predictors of high school students' academic resilience. *Journal of Educational Psychology, 93*(4), 758-771. https://doi.org/10.1037/0022-0663.93.4.758

Cassidy, S. (2016). The academic resilience scale (ARS-30): A new multidimensional construct measure. *Frontiers in Psychology*, *7*, 1787. https://doi.org/10.3389/fpsyg.2016.01787

Cheung, K. C. (2017). The effects of resilience in learning variables on mathematical literacy performance: A study of learning characteristics of the academic resilient and advantaged low achievers in Shanghai, Singapore, Hong Kong, Taiwan and Korea. *Educational Psychology, 37*(8), 965-982. https://doi.org/10.1080/01443410.2016.1194372

Cheung, K. C., Sit, P. S., Soh, K. C., Ieong, M. K., & Mak, S. K. (2014). Predicting academic resilience with reading engagement and demographic variables: Comparing Shanghai, Hong Kong, Korea, and Singapore from the PISA perspective. *The Asia-Pacific Education Researcher, 23*(4), 895-909. https://doi.org/10.1007/s40299-013-0143-4

Chirkina, T., Khavenson, T., Pinskaya, M., & Zvyagintsev, R. (2020). Factors of student resilience obtained from TIMSS and PISA longitudinal studies. *Issues in Educational Research, 30*(4), 1245-1263.

Collie, R. J., Ginns, P., Martin, A. J., & Papworth, B. (2017). Academic buoyancy mediates academic anxiety's effects on learning strategies: An investigation of English- and Chinese-speaking Australian students. *Educational Psychology, 37*(8), 947-964. http://dx.doi.org/10.1080/01443410.2017.1291910

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety, 18*(2), 76-82. https://doi.org/10.1002/da.10113

Cunningham, M., & Swanson, D. P. (2010). Educational resilience in African American adolescents. *Journal of Negro Education, 79*(4), 473-487.

Davidov, E., Muthen, B., & Schmidt, P. (2018). Measurement invariance in cross-national studies: Challenging traditional approaches and evaluating new ones. *Sociological Methods & Research, 47*(4), 631-636. https://doi.org/10.1177/0049124118789708

Desa, D. (2014). Evaluating measurement invariance of TALIS 2013 complex scales: Comparison between continuous and categorical multiple-group confirmatory factor analyses. *OECD Education Working Paper*, No. 103, OECD Publishing, Paris. https://doi.org/10.1787/5jz2kbbvlb7k-en

Erberer, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). Socioeconomically disadvantaged students who are academically successful: examining academic resilience cross-nationally. *IEA Policy Brief series, 5*.

Fantuzzo, J., Leboeuf, W., Rouse, H., Chen, C.-C., & Fantuzzo, J. (2012). Academic achievement of African American boys: A city-wide, community-based investigation of risk and resilience. *Journal of School Psychology, 50*(5), 559-579. https://doi.org/10.1016/j.jsp.2012.04.004

Fergus, S., & Zimmerman, M. A. (2005). Adolescent resilience: A framework for understanding healthy development in the face of risk. *Annu. Rev. Public Health, 26*, 399-419. https://doi.org/10.1146/annurev.publhealth.26.021304.144357

Findley, M. G., Kikuta, K., & Denly, M. (2021). External validity. *Annual Review of Political Science*, *24*, 365-393. https://doi.org/10.1146/annurev-polisci-041719-102556

Flannelly, K. J., Flannelly, L. T., & Jankowski, K. R. (2018). Threats to the internal validity of experimental and quasi-experimental research in healthcare. *Journal of Health Care Chaplaincy*, *24*(3), 107-130. https://doi.org/10.1080/08854726.2017.1421019

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist*, *18*(1), 12. https://doi.org/10.1027/1016-9040/a000124

Gabrielli, G., Longobardi, S., & Strozza, S. (2022). The academic resilience of native and immigrant-origin students in selected European countries. *Journal of Ethnic and Migration Studies, 48*(10), 2347-2368. https://doi.org/10.1080/1369183X.2021.1935657

Garcia-Crespo, F. J., Fernandez-Alonso, R., & Muniz, J. (2021). Academic resilience in European countries: The role of teachers, families, and student profiles. *Plos One, 16*(7). https://dx.doi.org/10.1371/journal.pone.0253409

Garcia-Crespo, F. J., Galian, B., Fernandez-Alonso, R., & Muniz, J. (2019). Educational resilience in reading comprehension: Determinant factors in PIRLS-Europe. *Revista de Educacion*, (384), 65-89. https://doi.org/10.4438/1988-592X-RE-2019-384-413

Garcia-Crespo, F. J., Suarez-Alvarez, J., Fernandez-Alonso, R., & Muniz, J. (2022). Academic resilience in mathematics and science: Europe TIMSS-2019 data. *Psicothema*, *34*(2), 217-225. https://doi.org/10.7334/psicothema2021.486

Garmezy, N., Masten, A. S., & Tellegen, A. (1984). The study of stress and competence in children: A building block for developmental psychopathology. *Child Development*, 97-111. https://doi.org/10.2307/1129837

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72. https://doi.org/10.1037/a0032138

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based Nursing*, *18*(3), 66-67. http://dx.doi.org/10.1136/eb-2015-102129

Hofmeyr, H. (2019). Performance beyond expectations: Academic resilience in South Africa, *Stellenbosch Economic Working Papers*, No. 19/2019, Retrieved from https://www.ekon.sun.ac.za/wpapers/2019/wp192019

Hughes, H. M., Graham-Bermann, S. A., & Gruber, G. (2001). Resilience in children exposed to domestic violence. In S. A. Graham-Bermann & J. L. Edleson (Eds.), *Domestic violence in the lives of children: The future of research, intervention, and social policy* (pp. 67-90). American Psychological Association. https://doi.org/10.1037/10408-004

Jang, E., Seo, Y. S., & Brutt-Griffler, J. (2023). Building academic resilience in literacy: digital reading practices and motivational and cognitive engagement. *Reading Research Quarterly*, *58*(1), 160-176. http://dx.doi.org/10.1002/rrq.486

Jin, S. L., Fang, G. B., Cheung, K. C., & Sit, P. S. (2022). Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.846466

Kane, M. T. (2006). Validity. In: R. L. Brennan (Eds.), *Educational Measurement* (4th edition, pp. 17-64), Westport, CT: American Council on Education & Praeger.

Kim, E. S., & Yoon, M. (2011). Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. *Structural Equation Modeling*, *18*(2), 212-228. https://doi.org/10.1080/10705511.2011.557337

Koirikivi, P., Benjamin, S., Hietajarvi, L., Kuusisto, A., & Gearon, L. (2021). Resourcing resilience: educational considerations for supporting well-being and preventing violent extremism amongst Finnish youth. *International Journal of Adolescence and Youth, 26*(1), 553-569. https://doi.org/10.1080/02673843.2021.2010578

Kosciw, J. G., Palmer, N. A., & Kull, R. M. (2015). Reflecting resiliency: Openness about sexual orientation and/or gender identity and its relationship to well-being and educational outcomes for LGBT students. *American Journal of Community Psychology, 55*(1-2), 167-178. https://doi.org/10.1007/s10464-014-9642-6

Lam, S. M., & Zhou, Y. (2021). SES-achievement gaps in East Asia: Evidence from PISA 2003-2018. *The Asia-Pacific Education Researcher, 31*, 691-710. https://doi.org/10.1007/s40299-021-00620-7

Langenkamp, A. G. (2010). Academic vulnerability and resilience during the transition to high school: The role of social relationships and district context. *Sociology of Education, 83*(1), 1-19. https://doi.org/10.1177/0038040709356563

Lansford, J. E., Malone, P. S., Stevens, K. I., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2006). Developmental trajectories of externalizing and internalizing behaviors: Factors underlying resilience in physically abused children. *Development and Psychopathology, 18*(1), 35-55. https://doi.org/10.10170S0954579406060032

Laukaityte, I. (2016). *Statistical modeling in international large-scale assessments.* Doctoral dissertation, Umeå Universitet.

Laukaityte, I., & Wiberg, M. (2017). Using plausible values in secondary analysis in large-scale assessments. *Communications in Statistics-Theory and Methods, 46*(22), 11341-11357. https://doi.org/10.1080/03610926.2016.1267764

Li, H., & Yeung, W. J. J. (2019). Academic resilience in rural Chinese children: Individual and contextual influences. *Social Indicators Research, 145*(2), 703-717. https://doi.org/10.1007/s11205-017-1757-3

Liew, J., Cao, Q., Hughes, J. N., & Deutz, M. H. (2018). Academic resilience despite early academic adversity: a three-wave longitudinal study on regulation-related resiliency, interpersonal relationships, and achievement in first to third grade. *Early Education and Development, 29*(5), 762-779. https://doi.org/10.1080/10409289.2018.1429766

Lines, R. L., Crane, M., Ducker, K. J., Ntoumanis, N., Thøgersen-Ntoumani, C., Fletcher, D., & Gucciardi, D. F. (2020). Profiles of adversity and resilience resources: A latent class analysis of two samples. *British Journal of Psychology, 111*(2), 174-199. https://doi.org/10.1111/bjop.12397

Luo, Y., Wang, A., Zeng, Y., & Zhang, J. (2022). A latent class analysis of resilience and its relationship with depressive symptoms in the parents of children with cancer. *Supportive Care in Cancer, 30*(5), 4379-4387. https://doi.org/10.1007/s00520-022-06860-7

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*, *71*(3), 543-562. https://doi.org/10.1111/1467-8624.00164

Luthar, S. S., & Zelazo, L. B. (2003). Research on resilience: An integrative review. *Resilience and vulnerability: Adaptation in the context of childhood adversities, 2*, 510-549. https://doi.org/10.1017/CBO9780511615788.023

Mäkikangas, A., Tolvanen, A., Aunola, K., Feldt, T., Mauno, S., & Kinnunen, U. (2018). Multilevel latent profile analysis with covariates: Identifying job characteristics profiles in hierarchical data as an example. *Organizational Research Methods*, *21*(4), 931-954. https://doi.org/10.1177/1094428118760690

Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., & Dicke, T. (2018). What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychological Methods*, *23*(3), 524. http://dx.doi.org/10.1037/met0000113

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, *47*(2), 106-124. https://doi.org/10.1080/00461520.2012.670488

Martin, A. J., Burns, E. C., Collie, R. J., Cutmore, M., MacLeod, S., & Donlevy, V. (2022). The role of engagement in immigrant students' academic resilience. *Learning and Instruction*, *82*, 1-9. https://dx.doi.org/10.1016/j.learninstruc.2022.101650

Martin, A. J., Ginns, P., Brackett, M. A., Malmberg, L.-E., & Hall, J. (2013). Academic buoyancy and psychological risk: Exploring reciprocal relationships. *Learning and Individual Differences*, *27*, 128-133. https://dx.doi.org/10.1016/j.lindif.2013.06.006

Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology*, *46*(1), 53-83. https://doi.org/10.1016/j.jsp.2007.01.002

Martin, M. O., Von Davier, M., & Mullis, I. V. (2020). *Methods and procedures: TIMSS 2019 technical report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/methods

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, *56*(3), 227. https://doi.org/10.1037/0003-066X.56.3.227

Masten, A. S., & Reed, M.-G. J. (2002). Resilience in development. *Handbook of Positive Psychology*, New York: The Guilford Press.

Mullis, I. V., & Martin, M. O. (2019). *PIRLS 2021 assessment frameworks*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/pirls2021/frameworks/

Mullis, I. V., & Martin, M. O. (2017). *TIMSS 2019 assessment frameworks*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2019/frameworks/

Munck, I., Barber, C., & Torney-Purta, J. (2018). Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009: The alignment method applied to IEA CIVED and ICCS. *Sociological Methods & Research*, *47*(4), 687-728. https://doi.org/10.1177/0049124117729691

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(6), 967-985. https://doi.org/10.1080/10705511.2019.1590146

OECD. (2011). *Against the odds: Disadvantaged students who succeed in school*. Paris: OECD Publishing. https://doi.org/10.1787/9789264090873-en

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264073234-en

OECD. (2019). *PISA 2018 Assessment and analytical framework*. PISA, OECD Publishing, Paris. https://doi.org/10.1787/b25efab8-en

Oldfield, J., Stevenson, A., & Ortiz, E. (2020). Promoting resilience in street connected young people in Guatemala: The role of psychological and educational protective factors. *Journal of Community Psychology*, *48*(2), 590-604. https://doi.org/10.1002/jcop.22272

Olsson, C. A., Bond, L., Burns, J. M., Vella-Brodrick, D. A., & Sawyer, S. M. (2003). Adolescent resilience: A concept analysis. *Journal of Adolescence*, *26*(1), 1-11. https://doi.org/10.1016/S0140-1971(02)00118-5

Özcan, B., & Bulus, M. (2022). Protective factors associated with academic resilience of adolescents in individualist and collectivist cultures: Evidence from PISA 2018 large scale assessment. *Current Psychology*, *41*(4), 1740-1756. https://doi.org/10.1007/s12144-022-02944-z

Pokropek, A., Davidov, E., & Schmidt, P. (2019). A monte carlo simulation study to assess the appropriateness of traditional and newer approaches to test for measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(5), 724-744. https://doi.org/10.1080/10705511.2018.1561293

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

Rudd, G., Meissel, K., & Meyer, F. (2021). Measuring academic resilience in quantitative research: A systematic review of the literature. *Educational Research Review*, *34*, 100402. https://doi.org/10.1016/j.edurev.2021.100402

Rutkowski, L., Gonzalez, E., Joncas, M., & Von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, *39*(2), 142-151. https://doi.org/10.3102/0013189X10363170

Rutkowski, D., & Rutkowski, L. (2013). Measuring socioeconomic background in PISA: One size might not fit all. *Research in Comparative and International Education*, *8*(3), 259-278. https://doi.org/10.2304/rcie.2013.8.3.259

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31-57. https://doi.org/10.1177/0013164413498257

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, *57*(3), 316- 331. https://doi.org/10.1111/j.1939-0025.1987.tb03541.x

Sacker, A., & Schoon, I. (2007). Educational resilience in later life: Resources and assets in adolescence and return to education after leaving school at age 16. *Social Science Research*, *36*(3), 873-896. https://doi.org/10.1016/j.ssresearch.2006.06.002

Sandoval-Hernández, A., & Bialowolski, P. (2016). Factors and conditions promoting academic resilience: A TIMSS-based analysis of five Asian education systems. *Asia Pacific Education Review*, *17*(3), 511-520. https://doi.org/10.1007/s12564-016-9447-4

Schoeneberger, J. A. (2016). The impact of sample size and other factors when estimating multilevel logistic models. *The Journal of Experimental Education*, *84*(2), 373-397. https://doi.org/10.1080/00220973.2015.1027805

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21-44). Boston, MA: Springer US.

Süleyman, A. (2022). Investigation of the individual characteristics that predict academic resilience. *International Journal of Contemporary Educational Research, 9*(3), 543-556. https://doi.org/10.33200/ijcer.1076091

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 111-130. https://doi.org/10.1080/10705511.2019.1602776

Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *International Journal of Academic Research in Management (IJARM), 5*. http://dx.doi.org/10.2139/ssrn.3205035

Teig, N., & Nilsen, T. (2022). Profiles of instructional quality in primary and secondary education: Patterns, predictors, and relations to student achievement and motivation in science. *Studies in Educational Evaluation, 74*, 101170. https://doi.org/10.1016/j.stueduc.2022.101170

Thiessen, V. (2008). Resilience and educational pathways: A longitudinal analysis of low reading achievers. *Canadian Journal of Family and Youth/Le Journal Canadien de Famille et de la Jeunesse*, 27-62. https://doi.org/10.29173/cjfy6047

Tudor, K. E., & Spray, C. M. (2017). Approaches to measuring academic resilience: A systematic review. *International Journal of Research Studies in Education, 7*(4). https://doi.org/10.5861/ijrse.2017.1880

Ungar, M. (2008). Resilience across cultures. *The British Journal of Social Work, 38*(2), 218-235. https://doi.org/10.1093/bjsw/bcl343

Ungar, M. (2011). The social ecology of resilience: Addressing contextual and cultural ambiguity of a nascent construct. *American Journal of Orthopsychiatry, 81*(1), 1. https://doi.org/10.1111/j.1939-0025.2010.01067.x

Ungar, M., & Liebenberg, L. (2011). Assessing resilience across cultures using mixed methods: Construction of the child and youth resilience measure. *Journal of Mixed Methods Research, 5*(2), 126-149. https://doi.org/10.1177/1558689811400607

Van De Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 486-492. https://doi.org/10.1080/17405629.2012.686740

Vicente, I., Pastor, J. M., & Soler, A. (2021). Improving educational resilience in the OECD countries: Two convergent paths. *Journal of Policy Modeling, 43*(6), 1149-1166. https://doi.org/10.1016/j.jpolmod.2021.09.007

Victor-Aigboidion, V., Onyishi, C. N., & Ngwoke, D. U. (2020). Predictive power of academic self-efficacy on academic resilience among secondary school students. *Journal of the Nigerian Council of Educational Psychologists, 12*(1).

Von Davier, M. (2020). Improving measurement properties of the PISA home possessions scale through partial invariance modeling. *Psychological Test and Assessment Modeling*, *62*(1), 55-83.

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI monograph series, 2*(1), 9-36.

Wang, M. C., Deng, Q., Bi, X., Ye, H., & Yang, W. (2017). Performance of the entropy as an index of classification accuracy in latent profile analysis: A Monte Carlo simulation study. *Acta Psychologica Sinica*, 49(11), 1473-1482. https://doi.org/10.3724/SP.J.1041.2017.01473

Wayman, J. C. (2002). The utility of educational resilience for studying degree attainment in school dropouts. *The Journal of Educational Research*, *95*(3), 167-178. https://doi.org/10.1080/00220670209596587

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: A guide to best practice. *Journal of Black Psychology*, *46*(4), 287-311. https://doi.org/10.1177/0095798420930932

Werner, E. E. (2000). Protective factors and individual resilience. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (pp. 115-132). Cambridge University Press. https://doi.org/10.1017/CBO9780511529320.008

Wills, G., & Hofmeyr, H. (2019). Academic resilience in challenging contexts: Evidence from township and rural primary schools in South Africa. *International Journal of Educational Research, 98*, 192-205. https://doi.org/10.1016/j.ijer.2019.08.001

Wu, Y. J., Chen, Y. H., Kiefer, S. M., & Carstensen, C. H. (2021). Learning strategies as moderators between motivation and mathematics performance in East Asian students: Latent class analysis. *Sage Open, 11*(4), 21582440211061382. https://doi.org/10.1177/21582440211061382

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49-54. https://doi.org/10.21315/eimj2019.11.2.6

Zolkoski, S. M., & Bullock, L. M. (2012). Resilience in children and youth: A review. *Children and Youth Services Review, 34*(12), 2295-2303. https://doi.org/10.1016/j.childyouth.2012.08.009

Appendices

Appendix A Risks, Positive Adaptations, and Scales Employed in Operationalizations

Measured by Components		Measured by Items
Risks	**Positive Adaptations**	
Simultaneous: assess risk and positive adaptation at the **same** time **Progressive**: assess risk and positive adaptation at **different** times		**Instrumental:** based on items
Risk Factors in Literature	*Protective Factors in Literature*	*Scales in Literature*
Demographic: race/ethnicity; immigrant status; ***Resources***: non-affluent schools; low- SES family; disadvantaged neighborhoods; rural area; ***Family***: problematic relationships with parents; single parent; with incarcerated parents; teen mother; foster care; unaccompanied refugee minors; ***Health or Life***: serious life difficulties; special needs students; preterm/low birth weight; sexual minority youth; ***Academic***: competitive academic environments; school-based discrimination; low academic achievement; dropout; struggle academically; extended school absences	***Academic***: performance; motivation, satisfaction, expectations; engagement; missed school days; academic behavior; degree attainment; high school completion; adult attainments; attend university; return to full-time education; not drop-out; ***Life***: well-being; managing lives well ; social inclusion;	Ego-resiliency scale (Block & Kremen, 1996) Child and Adolescent Functional Assessment Scale (CAFAS; Hodges,1999) Big Five Inventory (John & Srivastava, 1999) Motivation and engagement Scale High - School (MES-HS; Martin, 2002) Connor-Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003) Resilience Scale for Adults (Friborg, Hjemdal, Rosenvinge & Martinussen, 2003) Academic Resilience Inventory (ARI; Samouelz, 2004) General Academic Self-efficacy Scale (GASES; Abdul and Ashraf, 2006) Academic Buoyancy Scale (ABS; Martin & Marsh, 2008) Child and Youth Resilience Measure (CYRM-12; Liebenberg, Ungar & LeBlanc, 2013) Academic Risk and Resilience Scale (ARRS; Martin, 2013) Questionnaire for Anxiety and Resilience (De Beni et al., 2014) Academic Resilience Scale-30 (ARS-30, Cassidy 2016)

Appendix B Components for Home Resources for Learning (HRL), Home Educational Resources (HER), and Economic, Social and Cultural Status (ESCS)

Home Resources for Learning (HRL: TIMSS 2019/PIRLS2016 Grade 4)

Number of books in the home (students):

1) 0-10
2)11-25
3) 26-100
4) 101-200
5) More than 200

Number of home study supports (students):

1) None
2) Internet connection or own room
3) Both internet connection and own room

Highest level of education of either parent (parents):

1) Finished some primary or lower secondary or did not go to school
2) Finished lower secondary
3) Finished upper secondary
4) Finished post-secondary education
5) Finished university or higher

Highest level of occupation of either parent (parents):

1) Has never worked outside home for pay, general laborer, or semi-professional (skilled agricultural or fishery worker, craft or trade worker, plant or machine operator)
2) Clerical (clerk or service or sales worker)
3) Small business owner
4) Professional (corporate manager or senior official, professional, or technician or associate professional)

Number of children's books in the home (parents):

1) 0-10
2)11-25
3) 26-50
4) 50-100
5) More than 100

Home Educational Resources (HER: TIMSS 2019 Grade 8)

Number of books in the home:

1) 0-10
2)11-25
3) 26-100
4) 101-200
5) More than 200

Number of home study supports:

1) None
2) Internet connection or own room
3) Both internet connection and own room

Highest level of education of either parent:

1) Finished some primary or lower secondary or did not go to school
2) Finished lower secondary
3) Finished upper secondary
4) Finished post-secondary education
5) Finished university or higher

Economic, Social and Cultural Status (ESCS: PISA 2018)

Parents' highest level of education (PARED)

What is the <highest level of schooling> completed by your mother/father?

1) <ISCED level 3A>
2) <ISCED level 3B, 3C>
3) <ISCED level 2>
4) <ISCED level 1>
5) She did not complete <ISCED level 1>

Does your mother/father have this qualification? <ISCED level 6> (incl. higher qualifications at level 5A in some countries) (Yes or No)
Does your mother/father have this qualification? <ISCED level 5A> (excl. higher qualifications at level 5A in some countries) (Yes or No)
Does your mother/father have any of the following qualifications? <ISCED level 5B> (Yes or No)
Does your mother/father have any of the following qualifications? <ISCED level 4> (Yes or No)

Parents' highest occupational status (HISEI)

The following two questions concern your mother/father's job (open-ended questions):

What is your mother/father’s main job?
(e.g. school teacher, kitchen-hand, sales manager)
What does your mother/father do in his main job?
(e.g. teaches high school students, helps the cook prepare meals in a restaurant, manages a sales team)

Home possessions (HOMEPOS), including books in the home

Which of the following are in your home? (Yes or No)
A desk to study at
A room of your own
A quiet place to study
A computer you can use for school work
Educational software
A link to the Internet
Classic literature (e.g. <Shakespeare>)
Books of poetry
Works of art (e.g. paintings)
Books to help with your school work
<Technical reference books>
A dictionary
Books on art, music, or design
<Country-specific wealth item 1>
<Country-specific wealth item 2>
<Country-specific wealth item 3>

How many of these are there at your home? (Yes or No)
Televisions
Cars
Rooms with a bath or shower
<Cell phones> with Internet access (e.g. smartphones)
Computers (desktop computer, portable laptop, or notebook)
<Tablet computers> (e.g. <iPad®>, <BlackBerry® PlayBookTM>)
E-book readers (e.g. <KindleTM>, <Kobo>, <Bookeen>)
Musical instruments (e.g. guitar, piano)

How many books are there in your home?

1) 0-10
2) 11-25
3) 26-100
4) 101-200
5) 201-500
6) More than 500

Appendix C Missing Rates (%) of Home Resource for Learning (HRL) and Home Educational Resource (HER)

Country	T4 HRL	T8 HER	P4 HRL	Country	T4 HRL	T8 HER	P4 HRL
Albania	7.79	--	--	Lithuania	20.26	4.55	16.08
Argentina, Buenos Aires	--	--	33.80	Macao SAR	--	--	1.08
Armenia	4.48	--	--	Malaysia	--	0.45	--
Australia	100.00	1.74	56.16	Malta	28.79	--	13.49
Austria	10.91	--	6.56	Montenegro	14.80	--	--
Azerbaijan, Republic of	6.08	--	9.56	Morocco	10.73	0.89	15.43
Bahrain	10.48	0.72	11.84	Netherlands	100.00	--	46.60
Belgium (Flemish)	7.33	--	12.27	New Zealand	59.83	2.05	52.11
Belgium (French)	--	--	14.10	North Macedonia	14.98	--	--
Bosnia and Herzegovina	3.99	--	--	Northern Ireland	47.98	--	60.87
Bulgaria	2.60	--	1.75	Norway	43.28	6.67	6.81
Canada	38.12	--	19.36	Norway (4)	--	--	7.17
Canada (Ontario)	36.68	2.83	22.01	Oman	6.84	1.56	8.44
Canada (Quebec)	37.61	3.43	14.09	Pakistan	19.17	--	--
Chile	9.18	1.09	12.27	Philippines	6.47	--	--
Chinese Taipei	0.90	0.10	1.73	Poland	5.84	--	2.79
Croatia	1.96	--	6.05	Portugal	5.77	0.68	2.76
Cyprus	6.18	0.60	--	Qatar	21.63	1.26	19.53
Czech Republic	20.12	--	8.38	Romania	--	1.20	--
Denmark	42.39	--	7.37	Russian Federation	1.14	0.21	0.94
Egypt	--	2.15	--	Russian Federation, Moscow	1.28	0.21	1.66
England	100.00	5.41	100.00	Saudi Arabia	12.56	1.76	9.62
Finland	12.33	1.09	11.52	Serbia	3.29	--	--
France	12.09	3.87	5.17	Singapore	3.51	0.10	3.70
Georgia	9.82	1.54	32.61	Slovak Republic	4.33	--	4.42

Germany	40.24	--	3.64	Slovenia	--	--	5.40
Hong Kong, SAR	7.61	0.52	5.39	South Africa	23.43	0.99	--
Hungary	8.69	0.83	2.65	South Africa (Eng/Afr)	--	--	45.99
Iran, Islamic Republic of	2.70	0.17	7.66	South Africa (Gauteng)	--	0.64	--
Ireland	7.27	1.38	--	South Africa (Western Cape Province)	--	0.78	--
Israel	--	25.92	15.05	Spain	11.59	--	8.17
Italy	6.84	0.39	8.98	Spain (Andalucia)	--	--	8.83
Japan	3.53	0.18	--	Spain, Madrid	11.89	--	8.12
Jordan	--	1.37	--	Sweden	20.13	2.63	16.95
Kazakhstan	2.48	0.40	1.34	Trinidad And Tobago	--	--	15.71
Korea, Republic of	1.28	0.13	--	Turkey	6.93	0.81	--
Kosovo	4.20	--	--	United Arab Emirates	53.88	2.49	14.22
Kuwait	24.50	1.95	--	United Arab Emirates (Abu Dhabi)	63.88	4.33	18.89
Latvia	6.07	--	6.62	United Arab Emirates (Dubai)	50.31	0.28	13.58
Lebanon	--	1.33	--	United States	100.00	4.33	100.00

Note. T4 = TIMSS 2019 Grade 4; T8 = TIMSS 2019 Grade 8; P4 = PIRLS 2016 Grade 4; -- = Not Applicable, missing rate >= 20 is in red.

Part II

The Articles

1

2

3

1

Article 1

Ye, W., Teig, N., & Blömeke, S. (2023). Systematic review of protective factors related to academic resilience in children and adolescents: Unpacking the interplay of operationalization, data, and research method (under review in the journal of Educational Research Review)

Article 2

Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. *Educational Assessment, Evaluation and Accountability*, *33*(1), 169-202. https://doi.org/10.1007/s11092-020-09351-7

2

Educational Assessment, Evaluation and Accountability
https://doi.org/10.1007/s11092-020-09351-7

Academic resilience: underlying norms and validity of definitions

Wangqiong Ye, et al. ***[full author details at the end of the article]***

Received: 16 January 2020 / Accepted: 10 December 2020/Published online: 19 January 2021

Abstract
Academic resilience refers to students' capacity to perform highly despite a disadvantaged background. Although most studies using international large-scale assessment (ILSA) data defined academic resilience with two criteria, student background and achievement, their conceptualizations and operationalizations varied substantially. In a systematic review, we identified 20 ILSA studies applying different criteria, different approaches to setting thresholds (the same fixed ones across countries or relative country-specific ones), and different threshold levels. Our study on the validity of these differences and how they affected the composition of academically resilient students revealed that the classification depended heavily on the threshold applied. When a fixed background threshold was applied, the classification was likely to be affected by the developmental state of a country. This could result in an overestimation of the proportions of academically resilient students in some countries while an underestimation in others. Furthermore, compared to the application of a social or economic capital indication, applying a cultural capital indicator may lead to lower shares of disadvantaged students classified as academically resilient. The composition of academically resilient students varied significantly by gender and language depending on which indicator of human capital or which thresholds were applied reflecting underlying societal characteristics. Conclusions drawn from such different results depending on the specific conceptualizations and operationalizations would vary greatly. Finally, our study utilizing PISA 2015 data from three countries representing diverse cultures and performance levels revealed that a stronger sense of belonging to a school significantly increased the chances to be classified as academically resilient in Peru, but not in Norway or Hong Kong. In contrast, absence from school was significantly associated with academic resilience in Norway and Hong Kong, but not in Peru.

Keywords International large-scale assessments · Educational inequality · Human capital · Student achievement · Socio-economic background · Systematic review

1 Introduction

Resilience refers to successful adaption to situations despite risks that put someone at a disadvantage or adversity (Ungar 2005; Windle et al. 2011). In line with this general definition, academic resilience refers to the capacity of students to perform well in school despite a disadvantaged background (OECD 2011) or more precisely the heightened likelihood of success in school despite environmental adversities brought about by early traits, conditions, and experiences (Wang et al. 1994).

Since minimizing the influence of students' background on the outcomes of schooling is a central topic for accomplishing equity in education, a better understanding of academic resilience may help policymakers and educators to support students from a disadvantaged background in improving their academic performance. However, different conceptualizations of academic resilience may result in conflicting conclusions. It is therefore crucial to ensure the validity of a definition.

Studies on academic resilience typically employ some operationalization of socioeconomic status (SES) as an indicator of students' risk or adversity, and they use some type of educational outcome as an indicator of positive adaptation (Tudor and Spray 2017). Thresholds are usually used to combine continuous SES and outcome measures into a binary variable that indicates academic resilience or non-resilience.

In the context of international large-scale assessments (ILSAs), most studies adopted a composite SES index to operationalize students' background. General problems such as missing data or questionable comparability of this index across countries (Watermann et al. 2016), are specifically related to the conceptualization of academic resilience: A composite SES index treats student background as one-dimensional. Thus analyses based on such an index do not reveal the potential relevance of different SES components. Furthermore, studies applied different thresholds to define a disadvantaged background (and also to what it means to perform well). Whereas some studies used the same fixed thresholds for all countries included in their study, others used relative thresholds derived from the data within each country. The evidence supporting the validity of these decisions was often quite limited.

Since the measurement of academic resilience is inherently influenced by definitional issues, this study sought to examine the validity of different conceptualizations of academic resilience and how these affect the composition of academically resilient students. For this purpose, three countries were selected representing diverse cultures and performance levels (Norway, Peru, and Hong Kong). Student performance in science was used as an indicator of educational outcomes.

Besides the common composite SES index also used in other studies, three specific background indicators representing different dimensions of SES (economic, cultural, and social) were adopted to operationalize student background. Two types of thresholds (the same fixed and relative within-country thresholds) were applied to define a disadvantaged student background or high performance. Thus, in total, sixteen conceptualizations of academic resilience were examined on their validity, with four background indicators and two types of thresholds.

To illustrate how many and which students were classified as academically resilient, we selected two individual student characteristics (gender and language spoken at home). As validity measures, we selected two school-related characteristics (sense of belonging and absence from school) that can be supposed to assess similar concepts. This study examined their concurrent validity by comparing the relations of these external constructs to the different conceptualizations of academic resilience.

The paper is organized as follows. Firstly, a conceptual framework is developed that distinguishes between different ways to define academic resilience, including their underlying norms. Secondly, an overview of the literature about academic resilience is provided, in particular in the context of ILSAs. Research gaps and the research questions examined in this paper are presented thereafter. Thirdly, a methods section follows that provides information about the data and variables used and the analyses applied, results are presented after that. Finally, the paper concludes with a summary and a discussion of implications.

1.1 Conceptual framework: criteria of academic resilience and underlying norms

1.1.1 Resilience and academic resilience

Research on resilience in the behavioral sciences began to emerge around 1970. Since the mid-1980s, an increasing number of researchers from different disciplines (e.g., child development, pediatrics, psychology, psychiatry, and sociology) have published findings from studies on children who were successful in life despite adverse childhood environments (Werner 2000). The theoretical development about resilience has went through four waves: (1) identifying resilient qualities, (2) uncovering the resilience process, (3) promoting resilience through prevention and intervention, and (4) focusing on the dynamics of adaptation and change (Masten 2007). The latter means that resilience may vary across contexts and over time (Tudor and Spray 2017).

Although there is no universal definition for resilience across the different disciplines examining this phenomenon, most definitions are based around two core concepts: adversity and positive adaptation (Windle 2011). Correspondingly, in the context of schooling, academic resilience is defined by some measure of adversity in terms of early traits, conditions and experiences and by some measure of increased likelihood to succeed in school (Wang et al. 1994).

1.1.2 Measuring adversity: composite vs. distinct measures of student background

From a theoretical point of view, it is possible to distinguish between different dimensions (e.g., education, social status, and wealth) of an individual's background that may predefine his or her chances later in life. In major theories, the effects of social background on student outcomes are therefore conceptualized not only as a consequence of material possessions but also based on social and cultural practices (Bourdieu 1986). According to Bourdieu's capital theory (1986), individuals process economic, cultural, and social capital such as monetary resources, cultural possessions, and social relationships. These three types of capital can be distinguished, and each of them can be used for the accumulation of other types of capital.

Academic resilience studies typically use a composite index that covers several of these background dimensions. For example, the composite SES index of the PISA studies, economic, social, and cultural status (ESCS), includes parents' occupation, parents' education, and home resources (OECD 2017). Although the ESCS covers two of the three Bourdieu dimensions of capital, it is treated as a one-dimensional measure. Consequently, analyses conducted with this index cannot reveal the relevance of the different SES subdimensions for being academically resilient.

Studies on academic resilience using International Association for the Evaluation of Educational Achievement (IEA) data, for example TIMSS 2015, usually use the Home Educational Resources (HER) index, which is based on parents' education, the number of books at home, and home study support (Mullis and Martin 2013). It is therefore mostly a measure of students' cultural capital. Parents' occupation status as an indicator of students' economic capital was not included in the HER index, but was a part of another SES index for Grade four students, Home Resources for Learning (HRL).

Since the measurement of academic resilience is inherently influenced by conceptual issues (Windle et al. 2011), including alternative measures of social, cultural, and economic capital in the definition may shed light on how these dimensions of social background affect the results (Watermann et al. 2016). The present study follows this idea and assesses adversity with both composite and distinct measures.

A specific challenge is that the differences between countries make it challenging to use the same background measures to study academic resilience across countries (Coronado-Hijón 2017). For example, owning a car is often used as one indicator of student background, but this may have different meanings in economically developed and developing countries. Some of the measures used in the present study address this challenge; we will examine this issue further in the discussion.

1.1.3 Measuring positive adaptation: selecting an Indicator of student outcome

Educational outcomes can be distinguished into cognitive and non-cognitive (Heckman et al. 2006). Unlike resilience studies in psychology, non-cognitive outcomes were rarely used to measure positive adaptation in education (Tudor and Spray 2017). Non-cognitive skills like self-efficacy or educational aspiration were merely regarded as protective factors promoting academic resilience, or as outcomes of being resilient (OECD 2018). As a result, these studies tended to use cognitive outcomes, especially test scores to measure positive adaptation.

Test scores stem either from one or several subject domains. In case of using one subject domain, most studies focused on reading, mathematics, or science. These domains were regarded as providing fundamental skills needed for further education or success in the labor market (OECD 2018). Thus, one purpose of these studies was to shed light on the competitiveness of a country.

Considering that positive adaptation may vary by domain, some studies used data from different domains. OECD (2011) found students who showed positive adaptation in science did usually so also in mathematics or/and reading. However, other studies found that positive adaptation in one domain was not necessarily associated with positive adaptation in other domains. Therefore, they defined resilience as a characteristic across domains, for example, by showing positive adaptation in reading, mathematics, and sciences (Agasisti et al. 2018).

As previously stated, studies using cognitive skills to operationalize positive adaptation usually treated traits like anxiety, motivation, or engagement as predictors or outcomes associated with resilience due to bidirectional developmental processes (Coronado-Hijón 2017). Therefore, some educational researchers recently also began to use non-cognitive outcomes to assess positive adaptation in resilience studies (OECD 2018).

1.1.4 Thresholds for adversity and positive adaptation: cross-country vs. within-country

Despite decisions on selecting indicators of adversity and positive adaption, another step in conceptualizing academic resilience is to decide about the thresholds, which define a “disadvantaged” background (adversity) or “high” performance (positive adaptation). These decisions vary substantially across studies. One core distinction is between “fixed” and “relative” thresholds.“Fixed” means that the same threshold is applied across countries, whereas “relative” means that based on within-country data, different thresholds are used for different countries.

Using fixed thresholds stresses an international perspective where direct cross-country comparisons are at the forefront. In this perspective, the proportion of the resilient student is regarded as an indicator for quality and equity of education systems (Erberer et al. 2015; OECD 2011). Using relative thresholds means to define academic resilience from a national perspective, provides important insights on policy levers that are associated with resilience within different education systems (OECD 2011). When relative thresholds were applied, for example, successful disadvantaged students in one country may be classified as poor performing in other contexts.

A similar distinction as the one between fixed and relative thresholds is frequently made in the research on poverty that differentiates between absolute and relative poverty (Hagenaars and De Vos 1988). Research on academic resilience is more complex because it combines information from two criteria, student background and educational outcome. Therefore, we need to distinguish between four possible approaches to define academic resilience: (1) a fixed threshold for background and a fixed threshold for outcome; (2) a fixed threshold for background but a relative threshold for outcome; (3) a relative threshold for background but a fixed threshold for outcome; and (4) a relative threshold for background and a relative threshold for outcome. Several cutoff values (e.g. 20%, 25%, or 33%) were used to define thresholds in many studies; considering the economic and performance differences among our three samples, cutoff value 33% was adopted to have more students for analysis. Details are reported below.

1.2 State of research

1.2.1 Overview about academic resilience studies in international large-scale assessments

Since ILSAs have facilitated cross-country analyses of student achievement and its predictors, there is an increasing number of studies using data from ILSAs to investigate how individual and institutional features are related to academic resilience (Gonzalez and Padilla 1997; Martin and Marsh 2006; Sandoval-Hernández and Bialowolski 2016). To summarize the state of research, a systematic literature search in Web of Science, ERIC, and Google Scholar was carried out in July 2019. The search was built around four groups of key words: education (e.g., academic), resilience (e.g., resilient, buoyance), measurement (e.g., scale), and ILSA (e.g., PISA). The search was limited to English-language publications and revealed about 20 studies directly related to our topic (see Table 1). They applied a broad range of different criteria, different approaches to setting thresholds, and different threshold levels.

Table 1 shows the different operationalizations, which were grouped according to the four approaches to set thresholds explained above. As the overview reveals, most studies used Organization for Economic Co-operation and Development (OECD) data rather than IEA data. One reason could be the tremendous influence of PISA (Meyer

Table 1 Operationalizations of academic resilience in ILSAs

Type of threshold		Reference	Data	Student background indicator (threshold for defining adversity)	Outcome (threshold for defining positive adaptation)	Domain, Country samples	Main findings
Fixed background, fixed outcome	Direct	Erberer et al. (2015)	TIMSS 2011	HER index ("few resources")	Fix (≥ 475)	Mathematics (Grade 8), 28 education systems	Environments of high academic achievement appear to support academic resilience. Students' high educational aspirations appear to be the strongest and most consistent predictor of academic resilience.
		Sandoval-Hernández and Bialowolski (2016)	TIMSS 2011	HER index ("few resources")	Fix (≥ 475)	Mathematics (Grade 8), 5 Asian education systems	No consistent patterns were identified, student expectations and amount of time spent on homework were significantly associated with academic resilience in Singapore, boys tended to be more resilient in Korea than girls.
		Frempong et al. (2016)	TIMSS 2012	New SES index based on 18 assets listed in student questionnaire: disadvantaged students (bottom 2 SES quantiles) from disadvantaged schools (bottom 3 quantiles)	Fix (≥ 352 as the South African country mean)	Mathematics (Grade 9), South Africa	Girls who speak their native language tended to be more resilient. Positive attitude, valuing learning, job aspiration were important characteristics associated with academic resilience.
Fixed background, relative outcome		Sandoval-Hernandez and Cortés (2012)	PIRLS 2006	new SES index based on parents' education, occupation, home possessions (bottom 20% across countries in a cluster)	Within country (top 20%)	Reading, cluster of countries with comparable SES index	Non-cognitive characteristics (self-confidence and motivation) were the most important factors in predicting academic resilience. Material support provided by family and safety in school also mattered.
Relative background, fixed outcome		OECD (2011)	PISA 2009	ESCS (bottom 1/3 within country)	Across countries (top 1/3)	Science all participating countries	Academic resilience was not identified as a domain-specific characteristic. No gender gap in academic resilience was found for science. Language and

Table 1 (continued)

Type of threshold	Reference	Data	Student background indicator (threshold for defining adversity)	Outcome (threshold for defining positive adaptation)	Domain, Country samples	Main findings
						immigrant background were marginally associated with academic resilience in a few countries.
	OECD (2017)	PISA 2015 PISA 2006	ESCS (bottom 1/4 within country)	Across countries (top 1/4)	Science, all participating countries	In countries with a large increase in the proportion of academically resilient students, the percentage of low-performing students was reduced while the average performance was maintained or improved.
	OECD (2018)	PISA 2015	ESCS (bottom 1/4 within country)	Across countries (top 1/4)	Science, all participating countries	The presence of academic resilience varies greatly in this operationalization, 76% of disadvantaged students are academically resilient in Viet Nam, but less than 5% in Kosovo, Peru, and Tunisia.
	García-Crespo et al. (2019)	PIRLS 2016	New SES index based on possessions and books at home, highest academic qualifications, and level of employment of parents (bottom 25% within countries)	Across EU countries (top 25%)	Reading, EU countries that participated in this study	Student confidence in reading and a favorable school climate greatly increased the likelihood of academic resilience.
Residual	OECD (2010)	PISA 2009	ESCS (bottom 1/4 within country)	Across countries (top 1/4 of residual)	Reading, all participating countries	Academically resilient students were more prevalent in those education systems that PISA indicators showed to be more equitable.
	Cheung et al. (2014)	PISA 2009	ESCS (bottom 1/4 within country)	Across countries (top 1/4 of residual)	Reading, 4 East Asian economies	Family structure, expected education, kindergarten attendance, and reading engagement were associated with academic resilience.

Table 1 (continued)

Type of threshold	Reference	Data	Student background indicator (threshold for defining adversity)	Outcome (threshold for defining positive adaptation)	Domain, Country samples	Main findings
	Agasisti and Longobardi (2014)	PISA 2009	School ESCS and student SES (bottom 1/3 students from bottom 1/3 schools within country)	Across countries (top 1/3 of residual)	Reading, Italy	Motivation and extracurricular activities were positively associated with academic resilience. Immigrant status was negatively associated with academic resilience.
	Agasisti and Longobardi (2014)	PISA 2000–2012	School ESCS and student SES (bottom 1/3 students from bottom 1/3 schools within country)	Across countries (top 1/3 of residual)	Reading, 58 education systems over period 2000–2012	Educational funding can help disadvantaged students in overcoming their penalizing starting conditions.
	Cheung (2017)	PISA 2012	ESCS (bottom 1/4 within country)	Across countries (top 1/4 of residual)	Mathematics, Five East Asian education systems	Self-efficacy was strongly, familiarity with mathematical concepts learned in the earlier grades moderately, and anxiety weakly associated with academic resilience.
	Agasisti and Longobardi (2017)	PISA 2009	School ESCS and student SES (bottom 1/3 students from bottom 1/3 schools within country)	Across countries (top 1/3 of residual)	Reading, 5 EU countries	More extracurricular activities and a more positive school climate were positively associated with academic resilience.
Across-domain	Agasisti et al. (2018)	PISA 2006–2015	ESCS (bottom 1/4 within country)	≥ "Level 3" in all three domains	Reading, mathematics and science, all participating countries	Classroom disciplinary climate was positively associated with academic resilience. The amount of human and material resources available in school was weakly associated with academic resilience.
	OECD (2018)	PISA 2015	ESCS (bottom 1/4 within country)	≥ "Level 3" in all three domains	Reading, mathematics and science, all participating countries	This type of cross-domain academic resilience was more frequently observed in countries with higher average performance. Disciplinary climate and motivation were associated with academic resilience.

Table 1 (continued)

Type of threshold		Reference	Data	Student background indicator (threshold for defining adversity)	Outcome (threshold for defining positive adaptation)	Domain, Country samples	Main findings
	Non-cognitive	OECD (2018)	PISA 2015	ESCS (bottom 1/4 within country)	Socially and emotionally satisfied	Non-cognitive, 48 countries	Students who were socially and emotionally resilient also tended to do better academically. Lower shares of socio-emotionally resilient students were found in the top-performing systems than with the application of a cognitive outcome definition of resilience.
Relative background, relative outcome		OECD (2011)	PISA 2009	ESCS (bottom 1/3 within country)	Within country (top1/3)	Science, all participating countries	The extent to which students adopted positive approaches to learning and the amount of time they spent in regular science lessons were strongly associated with academic resilience.
		Karklina (2012)	PISA 2006 interview data	ESCS (bottom 1/3 within country)	Within country (top1/3)	Science, Latvia	Parental education level and length of poverty experience were associated with academic resilience.
		Aydiner and Kalender (2015)	PISA 2012	ESCS, bottom 1/4 within country	Within country (top1/3)	Reading, Turkey	Sense of belonging was associated with academic resilience
		OECD (2018)	PISA 2015	ESCS, bottom 1/4 within country	top1/4 within countries	Science, all participating countries	Academic resilience was not more frequently observed in top-performing countries than in developing or low-performing countries.

HER Home Educational Resources Scale; *ESCS* economic, social, and cultural status index; *School ESCS* average student ESCS in a school; *SES* socio-economic status; *Level 3* score higher than 484.14 and less than or equal to 558.73; *socially and emotionally satisfied* students satisfied with their life, felt socially integrated at school and did not suffer from test anxiety et al. 2017), another possible reason could be the missing data problem on student background indicators in IEA data (Broer et al. 2019). Therefore, studies using IEA data to explore academic resilience often either adopted a self-developed SES index (García-Crespo et al. 2019) or focused on selected countries with enough SES information (Cheung 2017; Erberer et al. 2015).

We will next review the ILSA research on academic resilience with respect to the conceptualizations and operationalizations used (see Table 1). For substantive results of these studies, please see the last column in this table.

1.2.2 Fixed background and fixed outcome thresholds

We identified three studies that used fixed thresholds to define both disadvantage and positive adaptation across different countries. Erberer et al. (2015) examined how prevalent academic resilience was across education systems and which protective factors could be identified. Their study used TIMSS 2011 data and adopted the composite Home Educational Resources (HER) index as a family SES measure. The authors classified a student as disadvantaged by applying a fixed threshold (a score ≤ 7.3 on the HER scale). Meanwhile, the authors used the so-called TIMSS International Intermediate Benchmark of Mathematics (students that reached this benchmark can apply basic mathematical knowledge in simple situations) as a threshold (a score ≥ 475) to define positive adaptation.

Sandoval-Hernández and Bialowolski (2016) adopted Erberer et al.'s (2015) method, and applied the definition to TIMSS 2011 data from five Asian education systems. Frempong et al. (2016) also followed this procedure and applied the definition to TIMSS 2011 data from South Africa. Frempong et al. did not adopt the HER index but calculated student SES index based on 18 assets listed in the student questionnaire.

In these three studies, the fixed thresholds for achievement to define positive adaptation were set either around the international (Erberer et al. 2015; Sandoval-Hernández and Bialowolski 2016) or the national mean (Frempong et al. 2016).

1.2.3 Fixed background and relative outcome thresholds

Our systematic review revealed only one study that adopted a fixed threshold to define adversity and a relative threshold to define positive adaptation. Sandoval-Hernández and Cortés (2012) applied the concept of academic resilience to Progress in International Reading Literacy Study (PIRLS) 2006 data. Since PIRLS 2006 does not provide a composite SES index (Mullis et al. 2004), authors followed Caro and Cortés's method (2012) and calculated an index based on parents' education, parents' occupation status, and home possessions. Considering measurement invariance, authors restricted their analysis to a cluster of countries with a comparable SES index. Disadvantaged background was defined by adopting a fixed SES threshold which was the 20th percentile of the index in the pooled data of all countries in the cluster. Positive adaptation was defined by a relative threshold which was the 80th percentile in each country.

1.2.4 Relative background and fixed outcome thresholds

Within this approach to define academic resilience, a methodological difference was found how to use the thresholds set. These were either used directly as in the studies described above or each disadvantaged student's performance was compared with the performance predicted by the average relationship among students from similar SES backgrounds across countries. The difference between these two was called a student's "residual" performance. Furthermore, within this group of studies, one of them used non-cognitive skills as an indicator of educational outcomes, and two of them used an across-domain operationalization of educational outcomes.

Direct threshold approaches OECD (2011) adopted the composite index ESCS and defined disadvantaged students by a relative background threshold (bottom 1/3 of ESCS within each country), whereas positive adaptation was defined by a fixed threshold (top 1/3 of students' performance across countries). OECD (2017) narrowed both thresholds down by defining academically resilient students as those who were in the bottom 1/4 of ESCS within each country and performed in the top 1/4 of students across all participating education systems. OECD (2018) adopted the same operationalization.

García-Crespo et al. (2019) explored predicting factors of academic resilience in reading literacy at Grade four, using PIRLS 2016 data from European Union member countries. The authors caculated their own Social, Economic, and Cultural Index (SECI) to measure student SES, based on home possession, number of books in the home, the highest academic qualifications of the parents, and the highest level of employment of the parents. Students in the bottom 25% of the SECI within each country, with a performance in the top 25% across the participating EU countries, were considered to be academically resilient.

Residual methods to calculate thresholds OECD (2010) defined disadvantaged students as those in the bottom 1/4 of ESCS within each country, while disadvantaged students in the top 1/4 of residual performance across countries were classified as academically resilient.

Several studies adopted this residual method, although OECD (2011) itself adopted new methods in its later studies (OECD 2018). Cheung et al. (2014) applied the residual method to PISA 2009 data from four East Asian economies in reading literacy. Academically resilient students were defined as those in the bottom 1/4 of ESCS within each country who achieved the top 1/4 residual performance across countries. Cheung (2017) applied the same definition to PISA 2012 data and examined academic resilience in mathematics, and also focused on a cluster of East Asian education systems.

Agasisti and Longobardi (2014, 2017) put special emphasis on disadvantaged students in disadvantaged schools and applied their definition to a group of European countries. The authors firstly selected schools among the 1/3 bottom of ESCS within each country based on the aggregated school ESCS average. From these schools, they selected those students who were in the 1/3 bottom of ESCS within the country. Resilient students were defined as disadvantaged students from disadvantaged schools who have a residual performance among the top 1/3 across countries.

Studies in this group usually focused on a cluster of countries with comparable economic and cultural backgrounds, because the strength between SES and performance varied across countries.

Across-domain operationalization of educational outcomes The studies mentioned above included only one domain (reading, science, or mathematics) as an indicator of positive adaptation. Agasisti et al. (2018) were the first to examine academic resilience across the three core domains in PISA—reading, mathematics, and science. Academically resilient students were defined as those among the bottom 1/4 of ESCS within each country, who performed at or above Proficiency Level 3 (i.e., one above the baseline level of proficiency needed to participate in society) in all three PISA domains. OECD (2018) adopted the same operationalization.

Outcome definition including non-cognitive characteristics Most studies in the ILSA context used cognitive outcomes (e.g., school achievement) to define positive adaptation, whereas non-cognitive skills (e.g., motivation) were treated as protective factors rather than indicators of positive adaptation. OECD (2018) examined for the first time non-cognitive outcomes and defined resilience in a non-cognitive way. Disadvantaged students from the bottom 1/4 of the ESCS distribution within each country were considered to be "socially and emotionally resilient", if they were satisfied with their life, felt socially integrated at school, and did not suffer from test anxiety (OECD 2018). When this definition was applied, lower shares of resilient students were found in the top-performing Asian educational systems than with the application of a cognitive outcome definition.

1.2.5 Relative background and relative outcome thresholds

Our systematic review identified four studies applying relative thresholds to defining both adversity and positive adaptations. As OECD (2011) mentioned, the purpose was to support policy makers and stakeholders with knowledge about how to foster resilience within their education systems. Disadvantaged students were defined by a relative threshold (bottom 1/3 of ESCS within each country), and the threshold for performance was also set as a relative one (top 1/3 within each country). Karklina (2012) used the same operationalization with PISA 2006 data from Latvia. Aydiner and Kalender (2015) adopted this approach as well but changed the cutoff values for thresholds—bottom 1/4 ESCS within a country for the background threshold and top 1/4 among the disadvantaged students within a country for the performance threshold. OECD (2018) followed this approach in its study about resilience from a national perspective and classified students from the bottom 1/4 of the ESCS distribution within each country and a performance among the top 1/4 of science within each country as resilient.

1.3 Does the conceptualization of academic resilience matter: a question of validity

In summary, validity may be defined as the extent to which we can back up the inferences drawn from an assessment by arguments based on evidence (Kane et al. 2005). The present study investigates the criterion validity of different conceptualizations of academic resilience, which means their relation to external criteria. Concurrent validity, where both the actual construct and the criterion measures are supposed to assess the same underlying trait and are collected at the same time, is a core dimension of criterion validity (Cohen and Swerdlik 2018).

Concurrent validity is demonstrated when a measure is positively or negatively correlated with another relevant measure as hypothesized, or when a new measure is associated with one that was already considered valid (Fink 2010). Two external criteria that should be strongly associated with academic resilience were applied, namely sense of belonging (positively) and absence from class (negatively), both of which have been identified in the literature as predictors of academic resilience (Sandoval-Hernández and Bialowolski 2016; Tommaso et al. 2018). A sense of belonging influences student outcomes via its effects on motivation and engagement, which were considered predictors of academic resilience in many studies (Aydiner and Kalender 2015; OECD 2011). Similarly, studies revealed that students who did not frequently skip class were more likely to be resilient (OECD 2018). The purpose of our study is to examine whether the strength of these relations varied by conceptualization of resilience.

Furthermore, we applied two background characteristics often used in the literature to describe the groups of students classified as academically resilient, namely gender and the language spoken at home (Cheung et al. 2014; OECD 2011) with the purpose to see how the group compositions changes depending on the conceptualization.

The aim of our study is to examine how different conceptualizations of academic resilience affect which students (gender and language) are classified as resilient, and to what extent the conceptualizations correspond with the two external criteria (sense of belonging and absence from school). Furthermore, given that the operationalization of student background may be affected by cultural differences and student performance varies substantially across countries, we examined concurrent validity for countries representing different cultures and performance levels.

2 The present study

As illustrated, four different approaches were applied to conceptualize academic resilience, either by using fixed or relative thresholds with respect to defining a disadvantaged student background or a strong educational outcome (positive adaptation). How the different approaches work empirically is largely an open question. For example, applying the same fixed thresholds to indicators of students' adversity may not work well in both developing and developed countries, because they may provide either very large or small groups of students classified as disadvantaged just because the whole country is less or more developed than others. Similarly, applying the same fixed threshold to indicators of positive adaptation may not work well in both high-performing and low-performing countries because they may lead to either very large or small groups of students classified as high-achieving just because the whole country performs more or less well. Furthermore, where thresholds were set varied substantially across studies. Little is known what these differences may mean regarding their relation to external criteria.

In addition, studies using PISA data to examine academic resilience can make use of a composite measure for SES (OECD 2005). However, the availability of this ESCS index is a double-edged sword to resilient studies because adopting the index without providing validity evidence or examining its subdimensions may underestimate their relevance (Watermann et al. 2016).

This study aims to examine the validity of different conceptualizations and operationalizations of academic resilience with data from countries that represent different developmental and achievement levels. Hong Kong was a high-achieving country (9th out of 72 participating countries), Norway was above average (24th), and Peru was near the bottom (66th) in the 2015 PISA cycle. Regarding the developmental status of these countries, Norway and Hong Kong both ranked highly on the human development index while Peru ranked low. It is worthwhile to mention that the rank of Hong Kong dropped dramatically in the inequality adjusted Human Development Index (UNDP 2016). Compared to other economies, Hong Kong has a relatively high-income disparity (Hong Kong Economy 2010). However, the relationship between SES and mathematics achievement was found to be the lowest among participating economies in PISA 2012 (Kalaycıoğlu 2015), which suggested high educational quality and equity in the system. Norway has a reputation for equity in its education system (Reimer et al. 2018), and empirical studies have found a pronounced increase of academic resilience in Norway from 2006 to 2015 (Agasisti et al. 2018). At the opposite extreme, empirical studies also found an extremely low percentage of resilient students in Peru. The three countries are in addition geographically separated and represent very different cultures.

Against this background, our study aims at answering the following questions:

1. How large is the group of academically resilient students when different conceptualizations of academic resilience are applied?
2. How do these conceptualizations of academic resilience affect which students are classified as academically resilient when it comes to gender and language background?
3. How are different conceptualizations of academic resilience associated with external variables, which can be supposed to assess similar constructs?
4. Do results change if different indicators of students' capital (economic, social, and cultural) are used?

3 Methods

3.1 Sample

This study used data from PISA 2015, which covered science, reading, mathematics, and financial literacy. Science was selected because it was the primary focus of this cycle and thus provided more precise estimates than for the other domains. Given the assumed relevance of a country's developmental and achievement state, this study used information from three education systems representing different economic contexts and performance levels—Hong Kong, Norway, and Peru.

PISA uses a two-stage stratified sampling strategy. Schools are sampled in the first stage, with the probability of selection being dependent on the number of eligible students enrolled. In the second stage, a random sample of students, aged from 15 years and 3 months to 16 years and 2 months, is selected within schools. Our total sample included 239 schools with an average about 24 students in Norway, 282 schools with an average 25 students in Peru, and 138 schools with an average 39 students in Hong Kong. Depending on the operationalization, the actual number of academically resilient students varied within countries from 137 to 5473 (for details see Appendix 1).

3.2 Measures

3.2.1 Disadvantaged student background: Adversity

PISA's composite SES index ESCS and three indicators of the SES subdimensions were used to examine how different conceptualizations of adversity affected the classification of students as academically resilient.

Economic, social, and cultural status The economic, social, and cultural status (ESCS) index is built on three components reflecting cultural and economic capital, thus two out of three of Bourdieu's dimensions of SES: parental education, parental occupation, and home possessions including books at home (OECD 2017). Social capital in the sense of Bourdieu (1986) is not covered by the ESCS.

The composite index is routinely used in resilience studies in the ILSA context; therefore, it was used in this study as well. The fixed threshold for ESCS was set to −0.68 across countries and reflected the bottom 1/3 of students internationally. The relative within-country thresholds were set to 0.25, −1.69, and −1.03 for Norway, Peru, and Hong Kong respectively and reflected the bottom 1/3 of the students nationally. Since the ESCS index was built on three standardized components via principal component analysis, some students may have the same ESCS score. For example, 25 students in Norway had the same ESCS score of 0.25 that represented the threshold. In this case, as many students were randomly selected out of the group of students with the same score as needed to end up with a group size of exactly 1/3. In the case of Norway, this meant to select 6 students (for details see Appendix 2).

Wealth As an index of students' economic capital, PISA provides an IRT scaled index called WEALTH, based on the number of material possessions. It includes 3 country-specific items and 9 items not directly related to educational support at home such as "Rooms with a bath or shower". The fixed threshold for WEALTH was set to −0.72 across countries reflecting the bottom 1/3 of all students internationally. The relative within-country thresholds were set to 0.26, −2.52, and −1.12 for Norway, Peru, and Hong Kong respectively.

Books This study used the variable "Number of books at home" (BOOKS) as an indicator of students' cultural capital. It consisted of 6 categories from 0 to 10 books to more than 500 books. The fixed threshold across countries was set to the second category (11–25 books), which included together with the first category a bit more than 1/3 of all students internationally. The relative within-country thresholds (bottom 1/3 of the students in each country) were set to the third category (26–100 books) for Norway, and to the second for both Peru and Hong Kong. In each country, several cases were in the category that reflected the threshold. To keep the number of disadvantaged students to the bottom 1/3 as intended, random cases were selected.

Parents' emotional support scale (EMOSUPS) Assessing students' social capital in ILSAs is challenging. Because it is typically represented by the relationship among family members that enhance the transmission of other resources (Bourdieu 1986), such as the family structure, parent-children discussion, parents' expectations and aspirations of children, parental education style, or intergeneration closure (Dika and Singh 2002). Since neither Norway nor Peru participated in the parent's questionnaire, which included several social capital indicators, our study used the parents' emotional support scale (EMOSUPS) from the student questionnaire as a proxy. The IRT scaled index was based on four statements such as whether students perceived their parents as supportive when they faced difficulties at school. The fixed threshold across countries was set to − 0.43, and the relative within country-thresholds were − 0.43, − .0.89, − 0.89 for Norway, Peru, and Hong Kong respectively. Random cases were selected if there were cases with the same score at the threshold.

To investigate whether classifications of resilience could be sensitive to the choice of the background indicator, we estimated Pearson correlations between them. The results indicated significant positive but imperfect associations among the four indicators (ESCS, WEALTH, BOOKS, and EMOSUPS). Basically, the composite index ESCS had strong or moderate associations with the indicators of economic (WEALTH) and cultural (BOOKS) capital, but only small associations with the indicator of social capital (EMOSUPS). Therefore, it is likely that the classification of disadvantaged students varies when different background indicators are applied to define student background.

3.2.2 Strong educational outcomes: positive adaptation

The outcome measure in this study is represented by the science score from PISA 2015. We used all 10 plausible values provided and combined the results of the ten separate analyses using Rubin's (1987) rules.

The fixed threshold was set at the mean of the PISA 2015 cycle of 466 points across countries. Disadvantaged students who scored higher than 466 were considered as resilient. The relative within-country thresholds, which represented the top 1/3 of the students within each education system, were set to 427.22 points for Peru, 542.61 points for Norway, and 561.13 points for Hong Kong.

3.2.3 Validity measures

Two student characteristics, gender and language spoken at home, were used to compare compositions of academically resilient students under different conceptualizations. Two external constructs, sense of belonging and absence from school, were adopted to examine the differences in strength between academic resilience and external constructs supposed to assess the same underlying idea across different definitions.

Gender Female students were coded as 1, male as 2. Gender is balanced in our sample. Further, 49.6% of the students were female in both Norway and Peru, and 49.5% were female in Hong Kong.

Language Students who usually used the same language as the language of the assessment were labeled as 1, and students who usually used another language were labeled as 2. About 8.7% and 7.2% of the students in Norway and Peru usually spoke another language at home; the proportion was smaller in Hong Kong, about 3.5%.

Sense of belonging Students were asked to rate six statements about their sense of belonging to school on a four-point Likert scale. We recoded the items so that higher scores referred to higher sense of belonging and built a latent variable BEL.

Absence from school PISA 2015 had three items to assess how often students skipped full school days, single classes, or arrived late at school. The four response categories ranged from "never" to "five or more times" with higher scores referring to higher absence from school. The three items were used to build the latent variable ABS.

Since BEL and ABS were both built on observed items assessed by a four-point Likert scale, we adopted Desa's (2014) suggestion and treated these items as categorical. A multiple-group confirmatory factor analysis (MG-CFA) was carried out to test measurement invariance of these constructs across Norway, Peru, and Hong Kong. This method usually relies on two-group comparison and is applied to relatively small sample sizes (Rutkowski and Svetina 2014). Therefore, comparisons across more than two groups with large sample sizes add complexity (Svetina et al. 2020). This study followed Svetina and Rutkowski's suggestion (2017) and adjusted the typical criteria to consider the changes in CFI greater than or equal to − 0.004 and changes in RMSEA less than or equal to 0.050 for evaluating metric invariance, and changes in CFI greater than or equal to − 0.004 and changes in RMSEA less than or equal to 0.010 for evaluating scalar invariance (Svetina et al. 2020). Analyses were conducted for BEL and ABS separately, metric invariance, and partial scalar invariance were established for both constructs (for details see Appendix 4).

3.3 Data analysis

The combination of four indicators of student background (1 composite SES index and 3 subdimensions of human capital) with two thresholds each time (1 fixed across countries and 1 relative within-country) led to eight operationalizations of adversity. In combination with the two versions of the indicator of positive adaption (science score with a fixed cross-country or a relative within-country threshold), we ended up with 16 different conceptualizations of academic resilience. The composition (gender and language spoke at home) of disadvantaged students were compared with the composition of the whole sample.

A logistic regression was fitted to estimate the likelihood of being resilient predicted by BEL and ABS. Data were handled and prepared in R (Version 3.6; R Core Team 2019), and analyses were conducted in Mplus (Version 8.4; Muthén and Muthén 2017). Missing data were handled by the default Full Information Maximum Likelihood (FIML) method in Mplus.

Students are nested in classes in our data set, but the nested data structure is not central to the research questions because all variables in question are located on the lowest level. Therefore, single-level models were estimated, which require fewer distributional assumptions and use a more parsimonious model approach (Stapleton et al. 2016). In the single-level logistic regression, student weights were used to make valid estimates and inferences of the population. Cluster characteristics due to the non-independence of samples (Stapleton et al. 2016) were taken into consideration by applying a sandwich estimator (type = complex in Mplus) and a robust weighted least squares estimator (WLSMV) so that standard errors and fit statistics were calculated properly (Asparouhov and Muthen 2006) (Fig. 1).

Ten plausible values combined with one background indicator led to 10 results, thus the resilient status of a student can be different in these 10 results. While these were combined in the estimation of the regression coefficients, results based on the first plausible value are presented to provide descriptive information about how the compositions of students change under different conceptualizations of academic resilience.

Two types of significant tests were applied. Firstly, a two sample *t* test was used to test the difference between proportions of disadvantaged students who are resilient when fixed and relative background thresholds were applied, or when fixed and relative performance thresholds were applied. The same type of *t* test was used to test differences between proportions of female and male students classified as academically resilient students or between students who speak the language of the test or another language. Secondly, a Wald chi-square test was used to examine whether coefficients of BLE and ABS differ across conceptualizations.

4 Results

4.1 How do different conceptualizations of academic resilience affect how much and which students are classified as academically resilient?

4.1.1 First step: defining disadvantaged students (adversity)

By implication, when relative thresholds for students' background indicators were applied, one out of three students in each country were classified as a disadvantaged student. Therefore, proportions of disadvantaged students were around 33.33% across conceptualizations using a relative background threshold. When the same fixed thresholds for

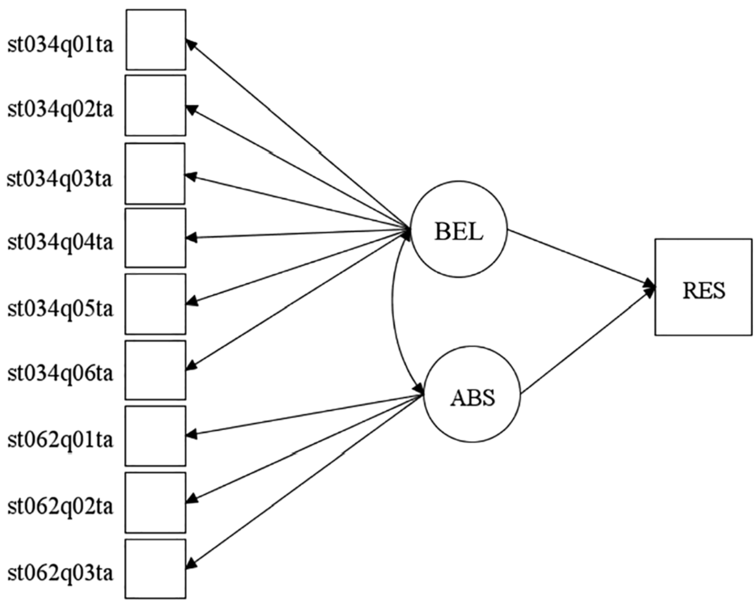


Fig. 1 Single-level logistic regression model. *RES* resilient, *BEL* sense of belonging, *ABS* absent from school, st034q01ta to st034q06ta are six observed items for sense of belonging, st062q01ta to st062q03ta are three observed items for absent from school.

background indicators were used across countries, results varied substantially (see Table 2). For example, if the fixed threshold of ESCS (− 0.68) was used to define student background, almost two-thirds of the students in Peru (63.94%) and half of the students in Hong Kong (46.76%) were considered as disadvantaged, but considerably fewer students in Norway (7.18%). Similar shifts were observed for the indicators of economic capital WEALTH and cultural capital BOOKS although on different levels.

In contrast, applying the background indicator of social capital, parents' emotional support (EMOSUPS) revealed a different pattern. While there were roughly one out of three students classified as disadvantaged in Norway (32.86%) and Peru (40.14%), the proportion was now highest in Hong Kong with almost two out of three students classified as disadvantaged (63.30%).

4.1.2 Second step: defining well-performing students (positive adaptation)

When relative within-country thresholds for science achievement were applied to define positive adaption, one out of three students was classified as well-performing in Norway, Peru, and Hong Kong. However, results varied substantially when the fixed threshold—set to the international PISA 2015 mean of 466 score points—was applied. Whereas in Norway almost two out of three students (63.54%) and in Hong Kong even more than three out of four students (78.17%) scored above the PISA 2015 mean, it was less than one out of five students in Peru (18.72%).

4.1.3 Third step: conceptualizing academically resilient students by combining background and performance definitions

There are in principle four possibilities to conceptualize academic resilience given what we have described above: combining a fixed or relative background threshold with the fixed performance threshold, or combining a fixed or relative background threshold with the relative performance threshold. We will systematically look at the results of these four approaches, firstly for the composite ESCS index and thereafter for the three subdimensions of human capital.

Combining the ESCS background thresholds with the fixed performance threshold When the fixed performance threshold of 466 points was combined with the ESCS indicator, about 70% of the disadvantaged students in Hong Kong and about half of the disadvantaged students in Norway were classified as being resilient, no

Table 2 Proportion of students classified as disadvantaged with fixed background thresholds

	Norway	Peru	Hong Kong
ESCS	7.18%	63.94%	46.76%
WEALTH	2.51%	78.51%	55.66%
BOOKS	20.80%	65.77%	36.41%
EMOSUPS	32.86%	40.14%	63.30%

ESCS index of economic, social, and cultural status; *EMOSUPS* index of parents' emotional support matter whether the fixed or the relative background thresholds were applied. Less than 10% of the disadvantaged students in Peru were classified as being academically resilient in both cases.

Given the low share of Norwegian students classified as disadvantaged with the fixed ESCS threshold applied across countries, the overall share of all students classified as academically resilient was very low (2.99%). The same low proportion of academically resilient students applied to Peru, but here because of the low share of students scoring above the PISA 2015 mean. In contrast, either one out of three or four Hong Kong students was classified as academically resilient (Table 3).

Combining the ESCS background thresholds with the relative performance thresholds When relative performance thresholds were applied, one out of three students in each country was defined as well performing. The proportions of disadvantaged students classified as academically resilient were more similar than in the cases described above with fixed performance thresholds across countries. The share varied only between 13.17% in Peru and 28.93% in Hong Kong, no matter whether the fixed or the relative background threshold was applied (see Table 4). Overall, this meant that the proportion of disadvantaged students classified as academically resilient went up in Peru and down in Norway and Hong Kong, when relative performance thresholds were applied instead of the same fixed performance threshold.

The same pattern occurred if we look at the proportion of all students classified as academically resilient. For example, with a fixed threshold of ESCS, the proportion of all students classified as academically resilient in Norway went down to 0.97% but went up in Peru to 14.55%.

Applying the subdimensions of human capital to define adversity The ESCS is a composite index that includes economic, cultural, and social capital indicators of student background at the same time. If we disentangle the conceptualization of adversity by applying the indicators of the three subdimensions separately, the data revealed similarities but also substantial differences. Depending on the type of thresholds applied either to background or to performance, the proportion of students classified as academically resilient varied.

In case of the same fixed performance threshold across countries, the pattern was similar. The proportion of disadvantaged students classified as academically resilient was highest in Hong Kong (between two thirds and three quarters of the disadvantaged students), followed

Table 3 Proportion of students classified as academically resilient when the fixed performance threshold was applied with the composite ESCS index

	Proportion of disadvantaged students classified as resilient			Proportion of all students classified as resilient		
	Norway	Peru	Hong Kong	Norway	Peru	Hong Kong
Fixed ESCS	41.58%	9.85%	73.18%	2.99%	6.30%	34.22%
Relative ESCS	50.06%	4.71%	70.71%	16.16%	1.56%	23.07%

ESCS index of economic, social, and cultural status

Table 4 Proportion of students classified as academically resilient when relative performance threshold were applied with the composite ESCS index

	Proportion of disadvantaged students classified as resilient			Proportion of all students classified as resilient		
	Norway	Peru	Hong Kong	Norway	Peru	Hong Kong
Fixed ESCS	13.52%	22.75%	28.93%	0.97%	14.55%	13.53%
Relative ESCS	20.89%	13.17%	26.66%	6.75%	4.37%	8.70%

ESCS index of economic, social, and cultural status

by Norway (between about 40% and 60% of the disadvantaged students) and lowest in Peru (between 5% and 17% of the disadvantaged students)—no matter whether the composite ESCS index or its subdimensions WEALTH, BOOKS, or EMOSUPS and no matter whether the fixed or the relative background thresholds were applied.

Although the pattern was similar, differences in the actual group size were visible resulting from variation in the proportion of students classified as disadvantaged (see Table 2). The indicator of economic (WEALTH) and in particular of social capital (EMOSUPS) led to higher shares of disadvantaged students classified as academically resilient than the indicator of cultural capital (BOOKS), no matter which type of background threshold was applied. This was particularly visible in Peru, where the proportion was up to twice or even three times as high as if an indicator of another subdimension or the composite ESCS index had been used (for details see Appendix 5).

The differences between applying the composite ESCS index or one of the subdimensions of human capital as indicators of adversity were even more pronounced when relative performance thresholds were used to define positive adaptation. The proportion of disadvantaged students classified as academically resilient was no longer highest in Hong Kong or lowest in Peru. For example, when relative performance thresholds were applied together with BOOKS as the indicator of cultural capital or EMOSUPS as the indicator of social capital, the proportion of disadvantaged students classified as academically resilient was lowest in Norway, no matter whether a fixed or a relative background threshold was used.

Similar to applying the fixed performance threshold, using WEALTH or EMOSUPS as capital indicators together with the relative performance threshold resulted usually in larger proportions of students classified as academically resilient than BOOKS (with one exception in Peru), regardless of whether a fixed or relative background threshold was used. It was particularly in Peru where the application of the indicator for social capital (EMOSUPS) increased the proportion of disadvantaged students classified as academically resilient (for details see Appendix 6).

4.2 Which students are classified as academically resilient in the different conceptualizations?

The proportions of students classified as academically resilient varied a lot by gender and language depending on the conceptualization and the country. Using the ESCS index as an indicator of adversity to define student background, there were between 31 female students (i.e., 0.44% of all students) in Peru and 905 female students (i.e., 16.89%) in Hong Kong classified as academically resilient, while the proportion of males varied between 0.37% (i.e., 20) in Norway and 17.34% (929) in Hong Kong (see Table 5). There were significantly fewer female than male students classified as academically resilient in Peru no matter which threshold was applied to define adversity or positive adaptation. If a relative outcome threshold was used, the same applied to Hong Kong. In contrast, Norway's proportions of female and male students classified as academically resilient were generally more balanced.

With respect to the language spoken at home in relation to the test language, there were between 2 students with a different language (i.e., 0.04% of all students) in Hong Kong and 97 students with a different language (i.e., 1.78%) in Norway classified as academically resilient, while the proportion of students with the same language classified as academically resilient varied between 0.88% (i.e., 48) in Norway and 33.74% (1808) in Hong Kong (see Table 6). In all three countries, there were significantly more students classified as academically resilient who spoke the test language at home, no matter which threshold was used to define adversity or positive adaptation.

Since the conceptualization of academic resilience included two criteria, namely student background with respect to adversity and student outcome with respect to positive adaptation, it was necessary to look at the criteria step-wise to be able to interpret the numbers presented above. We examined therefore the classification of students by gender and language firstly with respect to who was classified as disadvantaged, then at those who were classified as having high outcomes before we finally interpreted the combination in terms of academic resilience. As an overview, we started with applying ESCS as an indicator of student background, but we looked also at whether there were differences with respect to the economic, cultural, and social subdimensions of human capital.

Classification of academic resilience by gender There was a balanced gender distribution in Norway with respect to the classification of students as disadvantaged if the ESCS index was used as an indicator of adversity (see Appendix 9a). This result was

Table 5 Number of academically resilient students and percentage of all students by gender in each ountry

Back	Out	Norway				Peru				Hong Kong			
		Female		Male		Female		Male		Female		Male	
		n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Fix	Fix	94	1.72	69	1.26	159	2.28*	280	4.02	905	16.89	929	17.34
	Rel	33	0.60	20	0.37	427	6.13*	587	8.42	324	6.05*	401	7.48
Rel	Fix	463	8.49	419	7.68	31	0.44*	78	1.12	608	11.35	628	11.72
	Rel	190	3.48	178	3.26	120	1.72*	185	2.65	210	3.92*	256	4.78

Back student background (ESCS), *Out* student outcome (science achievement), *Fix* same fixed threshold across countries, *Rel* within-country threshold, * = proportion of female students significantly different from the proportion of males within the same operationalization, $p < 0.05$

Table 6 Number of academically resilient students and percentage of all students by language in each country

Back	Out	Norway				Peru				Hong Kong			
		Diff L		Same L		Diff L		Same L		Diff L		Same L	
		n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Fix	Fix	31	0.57*	132	2.42	7	0.10*	432	6.20	26	0.49*	1808	33.74
	Rel	5	0.09*	48	0.88	21	0.30*	993	14.24	2	0.04*	723	13.49
Rel	Fix	97	1.78*	785	14.39	4	0.06*	105	1.51	21	0.39*	1215	22.67
	Rel	33	0.60*	335	6.14	13	0.19*	292	4.19	2	0.04*	464	8.66

Back student background (ESCS), *Out* student outcome (science achievement), *Fix* same fixed threshold across countries, *Rel* within-country threshold, *L* language of the test spoken at home (different vs. same); * = the proportion of students who spoke another language at home was significantly different from the proportion of students who spoke test language at home in the same operationalization, $p < 0.05$

independent of the type of threshold applied. The proportions of female and male students in the top 1/3 Norwegian performers, i.e., in the group of those showing positive adaptation, were also evenly distributed—at least as long as the fixed performance threshold was used (see Appendix 11a). Applying the stricter relative performance threshold, there were significantly fewer females than males belonging to the top 1/3. As documented in Table 5, these performance differences were not large enough to affect the final gender distribution of academically resilient students, but they led to a significantly lower mean performance of the female students classified as academically resilient compared to males when the stricter relative within-country threshold in defining adversity was used (see Appendix 10a).

With respect to variation by subdimension of human capital (see Appendix 9a), the proportion of female students in Norway classified as disadvantaged was significantly lower than the proportion of males when BOOKS or Parents' emotional supports (EMOSUPS) were applied as indicators of adversity, independently of the type of threshold used. In contrast, significantly more female than male students were classified as disadvantaged, when the indicator WEALTH was applied together with the relative but not with the fixed threshold. The relative within-country threshold was much higher than applying the same fixed cross-country threshold.

Similar to Norway, if the ESCS was used as an indicator of adversity, there was a balanced gender distribution among disadvantaged students in Peru, independently of the type of threshold for adversity (see Appendix 9a). In Peru, the proportion of female students in the top 1/3 performers was significantly lower than of males, no matter which threshold indicating positive adaptation was applied (see Appendix 11a). As documented in Table 5, these performance differences led in turn to a significantly lower proportion of female than male students classified as academically resilient. Furthermore, the proportion of female students classified as academically resilient was significantly lower than the proportion of female students classified as disadvantaged in all operationalizations. In addition, the mean performance of academically resilient female students was often lower compared to male students, in particular when the more lenient relative within-country performance threshold was used (see Appendix 10a). Variation by subdimension of human capital was limited (see Appendix 9a).

Similar to Norway and Peru, there was no systematic difference in the gender distribution with respect to students' classification as disadvantaged using the ESCS index in Hong Kong (see Appendix 9a). The proportions of female and male students in the top 1/3 performers were evenly distributed if the lenient fixed performance threshold was used (see Appendix 11a). However, similar to Norway, applying the stricter relative performance threshold, there were significantly fewer females than males belonging to the top 1/3 (see Appendix 11a). Furthermore, the mean performance of academically resilient female students was often lower compared to male students (see Appendix 10a). With respect to the subdimensions of human capital, the same pattern was visible for Hong Kong as it had been documented for Norway when it came to BOOKS and EMOSUPS (see Appendix 9a). The proportion of female students classified as disadvantaged was significantly lower in these cases than of males, independently of the type of threshold used.

Classification of academic resilience by language Students who spoke a language at home different from the test language were significantly overrepresented in the group of students classified as disadvantaged with the ESCS index as an indicator of adversity in Norway (see Appendix 9b). This result was independent of the type of threshold applied. In contrast, they were significantly underrepresented in the top 1/3 Norwegian performers, i.e., in the group of those showing positive adaptation, again independently of the threshold used (see Appendix 11b). These differences strongly affected the distribution of academically resilient students by language (see Table 6). With respect to variation by subdimension of human capital (see Appendix 9b), when parents' emotional support (EMOSUPS) was applied as an indicator of adversity, independently of the type of threshold used, students who spoke another language at home were still significantly underrepresented but less dramatically as with the ESCS index.

The pattern was similar in Peru. Students who spoke a language at home different from the test language were significantly overrepresented in the group of students classified as disadvantaged no matter which threshold was used (see Appendix 9b). In contrast, they were significantly underrepresented in the top 1/3 performers, again independently of the threshold (see Appendix 11b). These differences strongly affected the distribution of academically resilient students by language (see Appendix 6). With respect to variation by subdimension of human capital (see Appendix 9b), the data revealed that using the WEALTH indicator together with the relative within-country threshold led to a stronger overrepresentation than with the other subdimensions.

The patterns were partly different in Hong Kong compared to Norway and Peru. There was no systematic effect of language on the classification as disadvantaged if the ESCS index was used (see Appendix 9b). However, students who spoke a different language at home were significantly underrepresented in the top 1/3 performers, independently of the threshold applied (see Appendix 11b). These performance differences strongly affected the distribution of academically resilient students by language (see Table 6). In regard to variation by subdimension of human capital (see Appendix 9b), the data revealed that using the BOOK indicator no matter which threshold was applied and using the relative WEALTH indicator led to a stronger overrepresentation of students with a different language than with the other subdimensions.

4.3 How are different conceptualizations associated with external variables?

In this section, we present the results of our study on concurrent validity including two external variables supposed to assess the same underlying idea. We present the relation between a classification of students as academically resilient and, firstly, their sense of belonging to a school and, secondly, their absence of school in terms of odds ratios based on standardized results (for details about model fit, estimates and 95% CI see Appendix 3).

Sense of belonging as the validity criterion Using the composite ESCS background indicator, sense of belonging (BEL) was statistically significantly and positively associated with academic resilience in all conceptualizations in Peru as hypothesized, but not in Norway or Hong Kong (see Table 7). These results indicate that students who said that they felt a higher sense of belonging to their school had a 30 to 40% higher chance to be classified as academically resilient in Peru. Unexpectedly, this did not apply to Norway or Hong Kong where no statistically significant effects of students' sense of belonging to their school on the classification as academically resilient was found.

With respect to the research question of differential results depending on the operationalization applied to academic resilience, there were no statistically significant differences for coefficients of BEL between fixed and relative performance thresholds or between fixed and relative background thresholds in case of using ESCS as an indicator of adversity. The same applied to using the number of books (BOOKS) or parents' emotional support (EMOSUPS) in all educational systems (see Appendix 7). However, there were statistically significant differences for coefficients of BEL between fixed and relative background thresholds in Peru when WEALTH was used.

Absence from school as the validity criterion Using the composite ESCS index as the background indicator, absence from school (ABS) was statistically significant and negatively associated with academic resilience in both Norway and Hong Kong as hypothesized (see Table 8). However in Peru, ABS was unexpectedly not significant in most operationalizations, and no noticeable pattern was found. A higher ABS score refers to higher frequency of absence from school; the results indicate that in Norway and Hong Kong, students who were more frequently absent from school had a 25 to 30% lower chance to be academically resilient.

Table 7 Odds ratio of effects of sense of belonging on the classification as academically resilient when ESCS was applied

		Norway ESCS		Peru ESCS		Hong Kong ESCS	
		Fixed	Relative	Fixed	Relative	Fixed	Relative
Performance	Fixed	1.05	1.00	1.30*	1.39*	0.95	0.97
	Relative	1.02	0.99	1.35*	1.40*	0.94	0.92

*$p < 0.05$, *ESCS* index of economic, social, and cultural status

Table 8 Odds ratios of effects of absence from school on the classification as academically resilient when ESCS was applied

		Norway		Peru		Hong Kong	
		ESCS		ESCS		ESCS	
		Fixed	Relative	Fixed	Relative	Relative	Relative
Performance	Fixed	0.72*	0.70*	0.94	1.00	0.69*	0.72*
	Relative	0.75*	0.71*	0.98	1.04	0.68*	0.69*

*$p < 0.05$, *ESCS* index of economic, social, and cultural status

With respect to the research question of differential results depending on the operationalization of academic resilience, there were no statistically significant differences for coefficients of ABS between fixed and relative performance thresholds or between fixed and relative background thresholds in Norway and Hong Kong, neither in case of ESCS nor in case of one of the subdimensions of human capital. This applied to Peru when the number of books (BOOKS) and parents' emotional support (EMOSUPS) were applied, but not for ESCS and WEALTH (see Appendix 8).

5 Discussion

This study examined the conceptualization and operationalization of academic resilience by firstly identifying adversity with indicators of student background, then by identifying positive adaptation with a performance indicator, and finally by combining these two criteria. In this process, many decisions had to be made: which indicator should be selected (e.g., composite or specific indicators in case of adversity), which type of threshold should be applied (the same fixed one across countries or relative within-country ones), which level should be set on each threshold (strict or lenient ones). The key finding of the present study was that the way how academic resilience was defined mattered. Different decisions resulted in different proportions of students who were classified as academically resilient and different compositions of the group of resilient students.

Our analyses revealed that fixed thresholds were frequently not well suited to identify academically resilient students across diverse countries. This applied to the background indicators as well as for the performance indicator, as the distributions of SES and performance measures varied considerably across economically developing and developed economies. For example, given the low share of Norwegian students classified as disadvantaged with the same fixed background threshold applied across countries, the overall share of all students classified as academically resilient was very low. The low proportion of academically resilient students was also found in Peru. However in this case, it was because of the low share of students scoring above the fixed performance threshold. In contrast, in Hong Kong, either one out of three or four students was classified as academically resilient in these cases. When a fixed threshold is used to define positive adaptation, the share of academically resilient students is heavily influenced by the level of economic development. As it was done for example by Sandoval-Hernández and Bialowolski (2016) or Erberer et al. (2015), a large share of academically resilient students does not necessarily equal to a better educational system as it was concluded in some reports (e.g., OECD 2011).

Generally, research on academic resilience is working with small sample sizes. This means we deal with strongly selected groups, which is not only a measurement challenge but raises policy questions as well. To what extent is it meaningful to implement activities in such a case? This challenge is increased by applying the same threshold across countries. For example, when a fixed ESCS threshold was combined with a relative performance threshold, only 53 Norwegian students were classified as academically resilient. Considering the whole sample size of Norway (5456), information about so few students is questionable if it shall be used to derive general measures beyond support of a specific group, which often is the conclusion in academic resilience papers.

SES refers to an individual or a family's position in a hierarchy according to access to wealth, power, and social status, and it usually includes parents' occupation, parents' education, and home income (Watermann et al. 2016). However, there were several concerns about these components. Researchers tended to use proxy to measure home income, but as mentioned before, owning a car means very different across countries. Measures about parents' occupation and education were ususlly collected from students' questionnaire, but students' responses may suffer from some degree of error, and the discordance varies across countries (Rutkowski and Rutkowski 2013). Therefore, comparisons of SES across country, for example, using a fixed SES threshold across countries, tend to raise measurement issues.

When relative within-country performance thresholds were applied to indicate positive adaptation, measurement problems with the student background indicators are no longer relevant. Furthermore, the proportions of disadvantaged students classified as academically resilient were more similar across countries than with the fixed performance threshold, because the proportions went up in Peru and down in Norway and Hong Kong. However, it seems worth to emphasize that the relative performance thresholds are hardly comparable in absolute terms, because they differed by 134 points (Peru 417 and Hong Kong 561), which corresponds to almost one and a half standard deviation on the PISA scale.

Based on Bourdieu's capital theory (1986), we adopted the three subdimensions of students' human capital separately to disentangle the effects of the composite ESCS conceptualization on academic resilience. The data revealed many similarities in the results but also substantial differences. The indicator of social capital led to higher shares of disadvantaged students classified as academically resilient than applying the composite index, and this particularly in Peru. This result may indicate a cultural difference between Peru and the other two systems, with social capital being educationally more relevant in Peru than in Norway and Hong Kong and thus more often revealed in the student survey in the first than in the latter systems. It would fit to results from cultural psychology where Peru often is characterized as a so-called collectivist country where social relations are more important than in so-called individualist countries such as Norway where people act rather independently (Hofstede and Peterson 2003).

The indicator of cultural capital resulted in all conceptualizations in smaller proportions of disadvantaged students classified as academically resilient than the composite ESCS index. This result may indicate a large spread in cultural capital than in the other types of human capital between disadvantaged and advantaged children. It seems to be harder for disadvantaged students to accomplish an amount of cultural capital that equals economic or social capital. The causal mechanism here is unknown though and should be examined in further research. It is worthwhile to point out that none of these differences resulting from applying one of the three subdimensions of human capital showed up when the composite ESCS index was used. The country differences in classifying disadvantaged students as academically resilient were often no longer significant. It seems as if advantages and disadvantages of the different approaches were balanced out in this case.

Concerning the composition of the group of students classified as academically resilient, the proportions varied by gender and language depending on the conceptualization and the country. This result points to the gender- and language-specific sensitivity of conceptualizations of academic resilience, and this in turn may reflect societal characteristics. It is therefore essential which operationalization of academic resilience is chosen in a study. The conclusions may be completely different. In many cases, it mattered for the composition of the group of academic resilient students whether the same fixed cross-country threshold was applied or a more or less lenient or strict relative within-country threshold. For example, applying the stricter relative performance threshold in Norway and Hong Kong led to significantly fewer females than males belonging to the top 1/3, which may be related to the general discussion about gender balance at the top of the performance distribution (Bergold et al. 2017). In Peru, the proportion of females was lower at the top no matter which threshold was applied.

It mattered also which subdimension of human capital was applied in several cases. The data revealed for example that both in Norway and Hong Kong the proportion of female students classified as disadvantaged was significantly lower than of males when the indicators of cultural or social capital were applied as indicators of adversity than the indicator of economic capital. Similarly, in two out of three countries, students who spoke a language at home different from the test language were less strongly underrepresented when the social capital indicator was applied, but more strongly when the economic capital indicator was applied. Educational inequalities between male and female students in student achievement, for example in terms of grades in mathematics or reading (Voyer and Voyer 2014), are well known for many countries and most researchers are aware of them. The same applies to educational inequalities depending on the language background of students (OECD 2018). Our study shows that it is important to pay attention to similar inequalities when it comes to background indicators as part of the definition of academic resilience as well.

When the same background indicator was applied with different performance thresholds, proportions of female or male students classified as academically resilient or students who spoke another language at home varied as well. If a fixed performance threshold was adopted, we were likely to underestimate the inequality in high-performing systems but overestimate the inequality in low-performing systems, thus exaggerate the inequality differences between high- and low-achieving systems. For language-related analysis, the data revealed that students who spoke another language at home were overrepresented in disadvantaged students but underrepresented in resilient students.

Finally, the results of our validity study revealed that the likelihood of being classified as resilient most often did not change significantly when different background thresholds or performance thresholds were applied. This result indicates some consistency in their definitions. At least the average size of relations to external constructs may not be affected too heavily.

6 Limitations

Before we turn to conclusions, it is necessary to point out some limitations of this study. Regarding indicators used in this study, student's performance in science was the only outcome indicator; analyses based on this domain may be not hold for other domains. Although we used four background indicators, ESCS as a composite SES index, WEALTH as an economic capital measure, BOOKS as a cultural capital measure, and EMOSUPS as a social capital index, all these indicators have limitations. For example, student's questionnaire of PISA 2015 had several items about social activities before and after school, including communications between parents and students, which would have been a good indicator of social capital. However, there was a large proportion of missing data (about 30%) from Peru. Therefore, this study adopted EMOSUPS as the only indicator to measure background from a social capital perspective. Finally, we have to deal with small groups in some cases. For example, when it comes to the research question which students were classified as academically resilient, the number of female students was as low as 15 students in Norway, and the number of students who spoke another language at home varied between 0 in Peru and 97 in Norway. The small numbers lead to large standard errors.

7 Conclusions

Proportions of resilient students are results of two criteria—a disadvantaged student background in terms of adversity and a high educational outcome in terms of positive adaptation. Our study shows how important it is to be careful with how to define both. How many and which students are classified as academically resilient is likely to be affected by the developmental state of a country, when a fixed threshold is applied to identify disadvantaged students. Thus, the pool of students who have a chance to be classified as academically resilient is predefined differently in each country. Countries that are highly developed economically may not have many disadvantaged students from a global perspective, which means that they by definition can have only very few academically resilient students. Considering the country-specific characteristics of background indicators, it may therefore be more meaningful to adopt a relative background threshold to define adversity.

In contrast, a decision on the type of performance threshold should largely depend on the aim of a study. If one would like to look at academic resilience across countries, for example with a global labor market in mind, a fixed performance threshold seems to be most meaningful. If one would like to look at academic resilience within a country, for example with the national labor market in mind, a relative performance threshold provides most information. However, given that some studies adopted the proportions of academically resilient students as an indicator for the quality and equity of an educational system (see, e.g., Agasisti et al. 2018), it is highly important also in this case to be aware of the changes in outcomes when different types of performance thresholds are applied. With a fixed performance threshold, we may conclude that Hong Kong has a higher level of quality and equity than Norway. If we replace the fixed performance thresholds with relative ones, proportions of resilient students in Norway and Hong Kong were no longer significantly different. Similarly, changes in the composition by gender and language have to be considered.

A conceptualization of academic resilience applying a relative background threshold and a fixed performance threshold was recommended by some studies (see, e.g., OECD 2011) to explore the presence of academic resilience across countries, or to study the consistency of individual and environmental characteristics associated with resilience. However, the shortcoming of this operationalization was also mentioned by OECD (2018). It may overestimate the amount of academic resilience in some countries while underestimate in some others. Since the share is closely related to students' average performance, higher-performing system tends to have bigger shares. However, it does not necessarily equal to a higher ability in helping disadvantaged students to exceed their predicted performance given their background, which is the essence of academic resilience. Although it may be too early to conclude that relative within-country thresholds work best also with respect to performance, these reflections point in any case to a strong need of more research that looks into the consequences of different conceptualizations of academic resilience. Researchers should feel encouraged to apply different approaches and to compare their results with respect to their robustness. Furthermore, they should not only pay attention to the overall proportion of students classified as academically resilient but also to their composition by gender or language.

Some conclusions can also be drawn regarding the choice of background indicators. The composite ESCS index as a multi-dimensional composite index seems to balance out to some extent different estimations happening when one of the subdimensions of human capital is applied. Whether this is appropriate is an open question and requires more research. The question here is, does an application of the subdimensions over- or underestimate the true size or do they indicate real differences.

Further research is thus needed in many respects. Research on academic resilience in education has deep roots in resilience studies carried out in psychology and sociology. However, there are differences we can learn from and which could move research on academic resilience forward. Firstly, unlike resilience studies in psychology, most studies in education are not longitudinal. It would be easier to identify causal mechanisms that increase academic resilience and may thereby contribute to overcoming educational inequality. Secondly, academic resilience studies have a methodological limitation in measuring adversity (Waxman et al. 2003), because disadvantaged students are treated as homogeneous groups despite possible variations in the degree to which their lives are actually affected by a risk (Luthar and Zelazo 2003). Our study has pointed to an approach to deal with this challenge by distinguishing between different types of human capital. Thirdly, most academic resilience studies in education treat non-cognitive skills as protective factors rather than including them as an indicator of positive adaptation. Acknowledging the relevance of, for example self-efficacy or educational aspiration as criterion for being academically resilience could change the discussion substantially.

References

Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). Academic resilience: What schools and countries do to help advantaged students succeed in PISA. OECD Education Working Papers, 167. https://doi.org/10.1787/e22490ac-en.

Agasisti, T., & Longobardi, S. (2014). Inequality in education: Can Italian disadvantaged students close the gap? *Journal of Behavioral and Experimental Economics, 52*(C), 8–20. https://doi.org/10.1016/j.socec.2014.05.002.

Agasisti, T., & Longobardi, S. (2017). Equality of educational opportunities, schools' characteristics and resilient students: An empirical study of EU-15 countries using OECD-PISA 2009 data. *An International and Interdisciplinary Journal for Quality-of-Life Measurement, 134*(3), 917–953. https://doi.org/10.1007/s11205-016-1464-5.

Asparouhov, T., & Muthen, B. (2006). Comparison of estimation methods for complex survey data analysis. *Mplus web notes*.

Aydiner, A., & Kalender, I. (2015). Student segments based on the factors related to sense of belonging across disadvantaged and resilient groups in PISA 2012. *Procedia - Social and Behavioral Sciences, 174*, 3299–3305. https://doi.org/10.1016/j.sbspro.2015.01.997.

Bergold, S., Wendt, H., Kasper, D., & Steinmayr, R. (2017). Academic competencies: Their interrelatedness and gender differences at their high end. *The Journal of Educational Psychology, 109*(3), 439–449. https://doi.org/10.1037/edu0000140.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Westport, CT: Greenwood Press.

Broer, M., Bai, Y., & Fonseca, F. (2019). *Socioeconomic inequality and educational*. Outcomes: Evidence from Twenty Years of TIMSS. https://doi.org/10.1007/978-3-030-11991-1.

Cheung, K. C. (2017). The effects of resilience in learning variables on mathematical literacy performance: A study of learning characteristics of the academic resilient and advantaged low achievers in Shanghai, Singapore, Hong Kong, Taiwan and Korea. *Educational Psychology, 37*(8), 965–982. https://doi.org/10.1080/01443410.2016.1194372.

Cheung, K. C., Sit, P. S., Soh, K. C., Ieong, M. K., & Mak, S. K. (2014). Predicting academic resilience with Reading engagement and demographic variables: Comparing Shanghai, Hong Kong, Korea, and Singapore from the PISA perspective. *The Asia-Pacific Education Researcher, 23*(4), 895–909. https://doi.org/10.1007/s40299-013-0143-4.

Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological testing and assessment: an introduction to tests and measurement (ninth edition. ed.)*. New York: McGraw-Hill Education.

Coronado-Hijón, A. (2017). Academic resilience: A transcultural perspective. *Procedia - Social and Behavioral Sciences, 237*(C), 594–598. https://doi.org/10.1016/j.sbspro.2017.02.013.

Desa, D. (2014). *Evaluating measurement invariance of TALIS 2013 complex scales*. https://doi.org/10.1787/5jz2kbbvlb7k-en.

Dika, S. L., & Singh, K. (2002). Applications of social Capital in Educational Literature: A critical synthesis. *Review of Educational Research, 72*(1), 31–60. https://doi.org/10.3102/00346543072001031.

Economy, H. K. (2010). Hong Kong's Gini coefficient compared with other economies. Retrieved from https://www.hkeconomy.gov.hk/en/pdf/gini_comparison.pdf:

Erberer, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-nationally. IEA Policy Brief series, 5. Retrieved from http://pub.iea.nl/fileadmin/user_upload/Policy_Briefs/IEA_Policy_Brief_Dec2013.pdf

Fink, A. (2010). Survey research methods. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., pp. 152–160). Oxford: Elsevier.

Frempong, G., Visser, M., Feza, N., Winnaar, L., & Nuamah, S. (2016). Resilient learners in schools serving poor communities. Electronic Journal of Research in Educational Psychology, 14(2), 352-367. https://doi.org/10.14204/ejrep.39.15038.

García-Crespo, F. J., Galián, B., Fernández-Alonso, R., & Muñiz, J. (2019). Educational resilience in reading comprehension: Determinant factors in PIRLS-Europe, Resiliencia Educativa en Comprensión Lectora: Factores determinantes en PIRLS-Europa. *Revista de Educacion, 2019*(384), 71–96. https://doi.org/10.4438/1988-592X-RE-2019-384-413.

Gonzalez, R., & Padilla, A. M. (1997). The academic resilience of Mexican American high school students. *Hispanic Journal of Behavioral Sciences, 19*(3), 301–317. https://doi.org/10.1177/07399863970193004.

Hagenaars, A., & De Vos, K. (1988). The definition and measurement of poverty. *The Journal of Human Resources, 23*(2), 211–221. https://doi.org/10.2307/145776.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482. https://doi.org/10.1086/504455.

Hofstede, G., & Peterson, M. (2003). Culture's consequences: Comparing values, behavior, institutions, and organizations across nations. In (Vol. 48, pp. 127-131).

Kalaycıoğlu, D. B. (2015). The influence of socioeconomic status, self-efficacy, and anxiety on mathematics achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. Educational Sciences: Theory and Practice, 15(5), 1391–1401. https://doi.org/10.12738/estp.2015.5.2731.

Kane, M., Crooks, T., & Cohen, A. (2005). Validating measures of performance. *Educational Measurement: Issues and Practice, 18*(2), 5–17. https://doi.org/10.1111/j.1745-3992.1999.tb00010.x.

Karklina, I. (2012). Family and children educational outcomes: social resilience within economically deprived families in Latvia. Paper presented at the rural environment. Education. Personality. (REEP). Proceedings of the international scientific conference Latvia.

Luthar, S. S., & Zelazo, L. B. (2003). Research on resilience: An integrative review. *Resilience and vulnerability: Adaptation in the context of childhood adversities, 2*, 510–549. https://doi.org/10.1017/CBO9780511615788.023.

Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*(3), 267–281. https://doi.org/10.1002/pits.20149.

Masten, A. S. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology, 19*(3), 921–930. https://doi.org/10.1017/S0954579407000442.

Meyer, H. D., Strietholt, R., & Epstein, D. Y. (2017). Three models of global education quality and the emerging democratic deficit in global education governance. In A. M. & L. G. (Eds.), International Handbook of Teacher Quality and Policy. New York: Routledge.

Mullis, I. V., Kennedy, A. M., Martin, M. O., & Sainsbury, M. (2004). PIRLS 2006 assessment framework and specifications: Progress in international reading literacy study. TIMSS & PIRLS, International Study Center, Lynch School of Education, Manressa House, Boston College, 140 commonwealth street, Chestnut Hill, MA 02467.

Mullis, I. V. S., & Martin, M. O. (2013). TIMSS 2015 assessment frameworks. In *Retrieved from Boston College*. TIMSS & PIRLS International Study: Center website http://timssandpirls.bc.edu/timss2015/frameworks.html.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

OECD. (2005). PISA 2003: Technical report. Retrieved http://www.pisa.oecd.org/dataoecd/49/60/35188570.pdf, from OECD.

OECD. (2010). *PISA 2009 results: Overcoming social background equity in learning opportunities and outcomes (volume II)*. Paris: OECD Publishing.

OECD. (2011). *Against the odds: Disadvantaged students who succeed in school*. Paris: OECD Publishing.

OECD. (2017). PISA 2015 technical report. Retrieved https://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final.pdf, from OECD.

OECD. (2018). *Equity in education: Breaking down barriers to social mobility*. Paris: PISA, OECD Publishing.

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

Reimer, D., Sortkear, B., Oskarsson, M., Nilsen, T., Rasmusson, M., & Nissinen, K. (2018). *Northern lights on TIMSS and PISA 2018*. Copenhagen: Nordisk Ministerråd.

Rutkowski, D., Rutkowski, L. (2013). Measuring Socioeconomic Background in PISA: One Size Might not Fit all. *Research in Comparative and International Education, 8*(3), 259–278

Rutkowski, L. & Svetina, D. (2014). "Assessing the hypothesis of measurement invariance in the context of large-scale international surveys." *Educational and Psychological Measurement, 74*(1), 31–57. https://doi.org/10.1177/0013164413498257.
Sandoval-Hernández, A., & Bialowolski, P. (2016). Factors and conditions promoting academic resilience: A TIMSS-based analysis of five Asian education systems. *Asia Pacific Education Review, 17*(3), 511–520. https://doi.org/10.1007/s12564-016-9447-4.
Sandoval-Hernandez, A., & Cortés, D. (2012). Factors and conditions that promote academic resilience: A cross-country perspective. In *Paper presented at the international congress for school effectiveness improvement (ICSEI)*. Malmö: Sweden.
Stapleton, L. M., McNeish, D. M., & Yang, J. S. (2016). Multilevel and single-level models for measured and latent variables when data are clustered. *Educational Psychologist, 51*(3–4), 317–330. https://doi.org/10.1080/00461520.2016.1207178.
Svetina, D., & Rutkowski, L. (2017). Multidimensional measurement invariance in an international context: Fit measure performance with many groups. *Journal of Cross-Cultural Psychology, 48*(7), 991–1008. https://doi.org/10.1177/0022022117717028.
Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 111–130. https://doi.org/10.1080/10705511.2019.1602776.
Tommaso, A., Francesco, A., Francesca, B., & Sergio, L. (2018). Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA: OECD publishing.
Tudor, K. E., & Spray, C. M. (2017). Approaches to measuring academic resilience: A systematic review. *International Journal of Research Studies in Education, 7*(4). https://doi.org/10.5861/ijrse.2017.1880.
UNDP. (2016). Human development report 2016: Human development for everyone. Retrieved from http://hdr.undp.org/sites/default/files/2016_human_development_report.pdf:
Ungar, M. (2005). Handbook for working with children and youth: Pathways to resilience across cultures and contexts.
Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174–1204. https://doi.org/10.1037/a0036620.
Wang, Haertal, G., & Walgberg, H. (1994). Educational resilience in inner-city. In M. Wang & G. E (Eds.), Educational resilience in inner-city America: Challenges and prospects (pp. 45–72). Hillsdale: Lawrence Erlbaunm Associates.
Watermann, R., Maaz, K., Bayer, S., & Roczen, N. (2016). Social background. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning: An international perspective* (pp. 117–145). Cham: Springer International Publishing.
Waxman, H. C., Gray, J. P., & Padron, Y. N. (2003). Review of research on educational resilience. In *eScholarship*. California: University of.
Werner, E. E. (2000). Protective factors and individual resilience. *Handbook of early childhood intervention, 2*, 115–132.
Windle, G. (2011). What is resilience? A review and concept analysis. *Reviews in Clinical Gerontology, 21*(2), 152–169. https://doi.org/10.1017/S0959259810000420.
Windle, G., Bennett, K. M., & Noyes, J. (2011). A methodological review of resilience measurement scales. *Health and Quality of Life Outcomes, 9*(1), 8–8. https://doi.org/10.1186/1477-7525-9-8.

Affiliations

Wangqiong Ye [1] **· Rolf Strietholt** [2] **· Sigrid Blömeke** [1]

✉ Sigrid Blömeke
sigrid.blomeke@cemo.uio.no

Wangqiong Ye
wangqiong.ye@cemo.uio.no

Rolf Strietholt
rolf.strietholt@tu-dortmund.de

[1] Centre for Educational Measurement, University of Oslo, Oslo, Norway

[2] Technische Universität Dortmund, Dortmund, Germany

Supplementary Material

Appendix 1

Size of the Different Subsamples Used in Our Studies Depending on the Conceptualization of Academic Resilience

Background	**Outcome**	**Norway**	**Subset**	**Peru**	**Subset**	**Hong Kong**	**Subset**
Fixed ESCS	Fixed Performance	1	392	1	4,457	1	2,506
	Relative Performance	2	392	2	4,457	2	2,506
Fixed WEALTH	Fixed Performance	3	137	3	5,473	3	2,983
	Relative Performance	4	137	4	5,473	4	2,983
Fixed BOOKS	Fixed Performance	5	1,135	5	4,585	5	1,951
	Relative Performance	6	1,135	6	4,585	6	1,951
Fixed EMOSUPS	Fixed Performance	7	1,793	7	2,798	7	3,392
	Relative Performance	8	1,793	8	2,798	8	3,392
Relative ESCS	Fixed Performance	9	1,762	9	2,315	9	1,748
	Relative Performance	10	1,762	10	2,315	10	1,748
Relative WEALTH	Fixed Performance	11	1,769	11	2,297	11	1,759
	Relative Performance	12	1,769	12	2,297	12	1,759
Relative BOOKS	Fixed Performance	13	1,759	13	2,304	13	1,758

	Relative Performance	14	1,759	14	2,304	14	1,758
Relative EMOSUPS	Fixed Performance	15	1,763	15	2,284	15	1,757
	Relative Performance	16	1,763	16	2,284	16	1,757

Note. ESCS=Index of Economic, Social and Cultural Status, EMOSUPS= Index of parents' emotional support.

Appendix 2

Cutoff Values for the Different Thresholds Applied

	Bottom 1/3 across all countries participating in PISA 2015				
Fix	**All**	**ESCS**	**WEALTH**	**BOOKS**	**EMOSUPS**
N of students	519,334	519,334	519,334	519,334	519,334
Missing (n)		15,963	15,730	18,046	92,004
missing (%)		3.07%	3.03%	3.47%	17.72%
valid N		503,371	503,604	501,288	427,330
Bottom 1/3	173,111	167,790	167,868	167,096	142,443
CUTOFF score		-0.68	-0.72	2	-0.43
n students with the same score		1,379	1,471	100,129	11,330
random pick to meet 1/3 criterion		236	638	68,401	5,706

	Bottom 1/3 within Norway				
Relative	Norway	ESCS	WEALTH	BOOKS	EMOSUPS
N	5456	5456	5456	5456	5456
Missing(n)		170	148	178	168
Missing(%)		3.12%	2.71%	3.26%	3.08%
Bottom 1/3	1819	1762	1769	1759	1763
cutoff score		0.25	0.26	3	-0.43
n students with the same score		25	41	1450	108

random pick to meet 1/3		6	3	623	60

	Bottom 1/3 within Peru				
Relative	Peru	ESCS	WEALTH	BOOKS	EMOSUPS
N	6971	6971	6971	6971	6971
Missing (n)		27	77	60	119
Missing(%)		0.39%	1.10%	0.86%	1.71%
Bottom 1/3	2324	2315	2298	2304	2284
cutoff score		-1.69	-2.52	2	-0.89
N students with the same score		22	0	2362	1107
random pick to meet 1/3		7	0	81	963

	Bottom 1/3 within Hong Kong				
Relative	Hong Kong	ESCS	WEALTH	BOOKS	EMOSUPS
N	5359	5359	5359	5359	5359
Missing(n)		115	82	85	87
missing(%)		2.15%	1.53%	1.59%	1.62%
bottom 1/3	1786	1748	1759	1758	1757
cutoff score		-1.03	-1.12	2	-0.89
n students with the same score		26	23	1067	1727

random pick to meet 1/3	18	6	876	287

Note. ESCS=Index of Economic, Social and Cultural Status, EMOSUPS= Index of parents' emotional support.

Appendix 3

Model fit, Estimates and 95% CI

					RES ON		L2.5%	U2.5%	L2.5%	U2.5%
					stdyx	stdyx	BEL		ABS	
	RMSEA	CFI	TLI	SRMR	BEL	ABS				
NOR1	0.037	0.993	0.989	0.038	0.047	-0.332	-0.137	0.232	-0.505	-0.158
NOR2	0.038	0.993	0.989	0.041	0.019	-0.292	-0.199	0.236	-0.557	-0.027
NOR3	0.102	0.947	0.917	0.080	0.010	-0.564	-0.355	0.376	-0.867	-0.262
NOR4	0.097	0.951	0.924	0.079	0.000	-0.405	-0.405	0.406	-0.805	-0.005
NOR5	0.037	0.994	0.991	0.031	0.051	-0.276	-0.073	0.175	-0.379	-0.173
NOR6	0.038	0.994	0.991	0.034	-0.018	-0.321	-0.178	0.142	-0.485	-0.158
NOR7	0.050	0.982	0.982	0.036	0.006	-0.320	-0.071	0.083	-0.395	-0.246
NOR8	0.049	0.989	0.982	0.036	-0.025	-0.296	-0.116	0.067	-0.382	-0.209
NOR9	0.044	0.992	0.988	0.034	-0.003	-0.353	-0.092	0.087	-0.441	-0.266
NOR10	0.044	0.992	0.988	0.037	-0.011	-0.343	-0.110	0.088	-0.454	-0.232
NOR11	0.042	0.992	0.988	0.032	0.017	-0.344	-0.066	0.100	-0.425	-0.262
NOR12	0.041	0.992	0.988	0.033	0.004	-0.327	-0.091	0.098	-0.429	-0.224
NOR13	0.039	0.993	0.990	0.030	0.014	-0.289	-0.072	0.100	-0.377	-0.202
NOR14	0.040	0.993	0.989	0.031	-0.017	-0.290	-0.132	0.098	-0.414	-0.166
NOR15	0.050	0.988	0.982	0.035	0.009	-0.318	-0.069	0.087	-0.394	-0.242
NOR16	0.049	0.989	0.982	0.035	-0.017	-0.285	-0.111	0.078	-0.373	-0.197
PER1	0.035	0.989	0.983	0.028	0.266	-0.057	0.188	0.344	-0.131	0.018
PER2	0.038	0.988	0.981	0.027	0.302	-0.022	0.242	0.361	-0.088	0.044

PER3	0.033	0.990	0.985	0.027	0.244	-0.070	0.181	0.307	-0.124	-0.017
PER4	0.035	0.989	0.983	0.026	0.295	-0.034	0.243	0.347	-0.090	0.023
PER5	0.032	0.991	0.986	0.026	0.260	-0.049	0.189	0.331	-0.119	0.021
PER6	0.036	0.989	0.982	0.027	0.289	-0.011	0.235	0.344	-0.081	0.058
PER7	0.041	0.983	0.973	0.032	0.240	-0.081	0.158	0.321	-0.157	-0.006
PER8	0.043	0.981	0.971	0.033	0.279	-0.050	0.215	0.344	-0.127	0.026
PER9	0.030	0.991	0.985	0.032	0.329	0.004	0.195	0.463	-0.145	0.153
PER10	0.032	0.990	0.984	0.029	0.339	0.044	0.249	0.428	-0.055	0.142
PER11	0.029	0.992	0.987	0.027	0.349	-0.011	0.230	0.468	-0.138	0.115
PER12	0.032	0.990	0.985	0.025	0.377	0.003	0.284	0.470	-0.088	0.093
PER13	0.039	0.986	0.978	0.033	0.293	-0.021	0.194	0.391	-0.126	0.084
PER14	0.043	0.983	0.974	0.031	0.310	0.033	0.225	0.394	-0.059	0.125
PER15	0.042	0.981	0.971	0.033	0.244	-0.076	0.154	0.335	-0.161	0.009
PER16	0.045	0.979	0.967	0.035	0.278	-0.032	0.203	0.353	-0.118	0.054
HKG1	0.032	0.991	0.986	0.032	-0.053	-0.370	-0.128	0.022	-0.472	-0.268
HKG2	0.032	0.991	0.986	0.033	-0.067	-0.385	-0.152	0.018	-0.521	-0.248
HKG3	0.034	0.988	0.982	0.039	-0.017	-0.353	-0.082	0.049	-0.443	-0.263
HKG4	0.036	0.987	0.980	0.040	-0.062	-0.333	-0.136	0.013	-0.443	-0.224
HKG5	0.035	0.990	0.985	0.031	-0.024	-0.292	-0.104	0.055	-0.394	-0.190
HKG6	0.035	0.990	0.984	0.031	-0.011	-0.294	-0.099	0.078	-0.463	-0.124
HKG7	0.030	0.992	0.988	0.033	0.002	-0.320	-0.059	0.064	-0.400	-0.239
HKG8	0.032	0.991	0.987	0.033	0.002	-0.306	-0.062	0.067	-0.398	-0.214
HKG9	0.033	0.992	0.987	0.034	-0.034	-0.323	-0.128	0.060	-0.455	-0.192
HKG10	0.033	0.992	0.987	0.035	-0.087	-0.373	-0.196	0.022	-0.551	-0.196

HKG11	0.040	0.988	0.982	0.045	0.018	-0.285	-0.062	0.098	-0.402	-0.169
HKG12	0.041	0.988	0.981	0.047	-0.033	-0.327	-0.136	0.069	-0.474	-0.181
HKG13	0.033	0.990	0.985	0.032	-0.039	-0.289	-0.125	0.046	-0.395	-0.184
HKG14	0.034	0.990	0.985	0.033	-0.039	-0.312	-0.138	0.060	-0.492	-0.131
HKG15	0.037	0.987	0.979	0.039	-0.004	-0.350	-0.094	0.085	-0.453	-0.246
HKG16	0.036	0.987	0.979	0.039	-0.010	-0.327	-0.100	0.081	-0.449	-0.205

Note. BEL= sense of belonging, ABS=absence from school, RES= resilient, CFI= Comparative fit index, RMSEA= root mean square error of approximation.

Appendix 4

Fit Indices Results of Measurement Invariance Test

BEL	χ^2	df	p value	CFI	RMSEA	Δ CFI	Δ RMSEA
CONFIGURAL MODEL	53.57	3	0	0.999	0.055		
METRIC MODEL	207.89	15	0	0.998	0.048	-0.001	-0.007
SCALAR MODEL (partial)	17.722	5	0.003	0.999	0.021	0.001	-0.027

ABS	χ^2	df	p value	CFI	RMSEA	Δ CFI	Δ RMSEA
CONFIGURAL MODEL	0	0	NA	1	0		
METRIC MODEL	6.86	6	0.334	1	0.005	0	0.005
SCALAR MODEL (partial)	24.09	8	0.002	0.999	0.013	-0.001	0.008

Note. BEL= sense of belonging, ABS=absence from school, RES= resilient, df=degree of freedom, CFI= Comparative fit index, RMSEA= root mean square error of approximation, partial scalar invariance for BEL was established by releasing two loadings (ST034Q01TA and ST034Q06TA), partial scalar invariance for ABS was established by releasing one loading (ST062Q23TA).

Appendix 5

Proportion of Disadvantaged Students Classified as Academically Resilient When the Fixed Performance Threshold Was Applied With WEALTH, BOOKS and EMOSUPS as Background Indicators

	Norway	**Peru**	**Hong Kong**
Fixed WEALTH	56.20%	14.00%	76.47%
Fixed BOOKS	39.30%	11.52%	67.45%
Fixed EMOSUPS	58.51%	17.33%	76.24%
Relative WEALTH	61.90%	5.35%	73.74%
Relative BOOKS	47.13%	7.60%	66.78%
Relative EMOSUPS	58.37%	17.43%	74.90%

Note. EMOSUPS= Index of Parents' Emotional Support.

Appendix 6

Proportion of Disadvantaged Students Classified as Academically Resilient When the Relative Performance Thresholds Were Applied With WEALTH, BOOKS and EMOSUPS

	Norway	**Peru**	**Hong Kong**
Fixed WEALTH	23.36%	28.16%	32.08%
Fixed BOOKS	12.78%	24.62%	23.01%
Fixed EMOSUPS	27.11%	32.67%	32.99%
Relative WEALTH	31.26%	14.24%	30.53%
Relative BOOKS	17.85%	18.66%	23.38%
Relative EMOSUPS	26.77%	32.66%	31.99%

Note. EMOSUPS= Index of Parents' Emotional Support.

Appendix 7

Coefficient Comparisons for Sense of Belonging

Coefficient Comparisons for BEL Between Operationalizations With Same Background Threshold but Different Performance Thresholds

BEL Fixed Performance		Relative Performance		chisq	df	p	SS
NOR1	0.058	NOR2	0.023	0.090	1	0.764	
NOR3	0.013	NOR4	0.000	0.003	1	0.957	
NOR5	0.061	NOR6	-0.022	1.190	1	0.275	
NOR7	0.008	NOR8	-0.025	0.435	1	0.510	
NOR9	-0.003	NOR10	-0.013	0.032	1	0.857	
NOR11	0.020	NOR12	0.004	0.101	1	0.751	
NOR13	0.016	NOR14	-0.020	0.498	1	0.480	
NOR15	0.011	NOR16	-0.021	0.423	1	0.516	
PER1	0.323	PER2	0.367	0.812	1	0.367	
PER3	0.300	PER4	0.363	2.553	1	0.110	
PER5	0.317	PER6	0.353	0.668	1	0.414	
PER7	0.290	PER8	0.339	0.961	1	0.327	
PER9	0.417	PER10	0.429	0.020	1	0.887	
PER11	0.428	PER12	0.464	0.238	1	0.625	
PER13	0.361	PER14	0.382	0.109	1	0.741	
PER15	0.294	PER16	0.336	0.565	1	0.452	
HKG1	-0.066	HKG2	-0.084	0.144	1	0.704	
HKG3	-0.021	HKG4	-0.079	1.812	1	0.178	

HKG5	-0.031	HKG6	-0.014	0.108	1	0.743	
HKG7	0.003	HKG8	0.003	0.000	1	0.999	
HKG9	-0.043	HKG10	-0.109	1.211	1	0.271	
HKG11	0.023	HKG12	-0.042	1.569	1	0.210	
HKG13	-0.050	HKG14	-0.050	0.000	1	1.000	
HKG15	-0.006	HKG16	-0.013	0.013	1	0.910	

Coefficients Comparison for BEL Between Operationalizations With Same Performance Threshold but Different Background Thresholds

BEL							
Fixed Performance		Relative Performance		chisq	df	p	SS
NOR1	0.058	NOR9	-0.003	0.277	1	0.599	
NOR2	0.023	NOR10	-0.013	0.070	1	0.792	
NOR3	0.013	NOR11	0.020	0.001	1	0.979	
NOR4	0.000	NOR12	0.004	0.000	1	0.989	
NOR5	0.061	NOR13	0.016	0.349	1	0.555	
NOR6	-0.022	NOR14	-0.020	0.000	1	0.987	
NOR7	0.008	NOR15	0.011	0.005	1	0.946	
NOR8	-0.025	NOR16	-0.021	0.030	1	0.864	
PER1	0.323	PER9	0.417	3.722	1	0.054	
PER2	0.367	PER10	0.429	2.767	1	0.096	
PER3	0.300	PER11	0.428	10.506	1	0.001	***
PER4	0.363	PER12	0.464	9.422	1	0.002	***
PER5	0.317	PER13	0.361	0.995	1	0.318	

PER6	0.353	PER14	0.382	0.735	1	0.391
PER7	0.290	PER15	0.294	0.006	1	0.937
PER8	0.339	PER16	0.336	0.007	1	0.935
HKG1	-0.066	HKG9	-0.043	0.238	1	0.626
HKG2	-0.084	HKG10	-0.109	0.221	1	0.638
HKG3	-0.021	HKG11	0.023	1.084	1	0.298
HKG4	-0.079	HKG12	-0.042	0.569	1	0.451
HKG5	-0.031	HKG13	-0.050	0.134	1	0.715
HKG6	-0.014	HKG14	-0.050	0.396	1	0.529
HKG7	0.003	HKG15	-0.006	0.048	1	0.827
HKG8	0.003	HKG16	-0.013	0.135	1	0.713

Note. BEL= sense of belonging, NOR= Norway, PER= Peru, HKG= Hong Kong, SS= statistically significant.

Appendix 8

Coefficient Comparisons for Absence From School

Coefficients Comparison for ABS Between Operationalizations With Same Background Threshold but Different Performance Thresholds

ABS							
Fixed Performance		Relative Performance		chisq	df	p	SS
NOR1	-0.412	NOR2	-0.371	0.139	1	0.709	
NOR3	-0.641	NOR4	-0.459	0.958	1	0.328	
NOR5	-0.309	NOR6	-0.360	0.761	1	0.383	
NOR7	-0.371	NOR8	-0.296	2.754	1	0.097	
NOR9	-0.404	NOR10	-0.397	0.018	1	0.893	
NOR11	-0.412	NOR12	-0.395	0.117	1	0.732	
NOR13	-0.325	NOR14	-0.327	0.001	1	0.976	
NOR15	-0.368	NOR16	-0.333	0.594	1	0.441	
PER1	-0.078	PER2	-0.031	0.789	1	0.374	
PER3	-0.094	PER4	-0.045	1.771	1	0.183	
PER5	-0.066	PER6	-0.015	1.100	1	0.294	
PER7	-0.116	PER8	-0.072	0.644	1	0.422	
PER9	0.006	PER10	0.061	0.272	1	0.602	
PER11	-0.015	PER12	0.004	0.048	1	0.828	
PER13	-0.028	PER14	0.043	0.998	1	0.318	
PER15	-0.108	PER16	-0.046	1.014	1	0.314	
HKG1	-0.411	HKG2	-0.433	0.122	1	0.727	
HKG3	-0.399	HKG4	-0.377	0.155	1	0.694	

HKG5	-0.328	HKG6	-0.330	0.002	1	0.966	
HKG7	-0.343	HKG8	-0.329	0.093	1	0.760	
HKG9	-0.351	HKG10	-0.403	0.511	1	0.475	
HKG11	-0.324	HKG12	-0.369	0.437	1	0.508	
HKG13	-0.326	HKG14	-0.350	0.148	1	0.701	
HKG15	-0.382	HKG16	-0.355	0.204	1	0.652	

Coefficients Comparison for ABS Between Operationalizations With Same Performance Threshold but Different Background Thresholds

ABS				chisq	df	p	SS
Fixed Performance		Relative Performance					
NOR1	-0.412	NOR9	-0.404	0.005	1	0.942	
NOR2	-0.371	NOR10	-0.397	0.023	1	0.879	
NOR3	-0.641	NOR11	-0.412	1.518	1	0.218	
NOR4	-0.459	NOR12	-0.395	0.072	1	0.789	
NOR5	-0.309	NOR13	-0.325	0.078	1	0.780	
NOR6	-0.360	NOR14	-0.327	0.124	1	0.725	
NOR7	-0.371	NOR15	-0.368	0.003	1	0.953	
NOR8	-0.296	NOR16	-0.333	0.048	1	0.826	
PER1	-0.078	PER9	0.006	2.542	1	0.111	
PER2	-0.031	PER10	0.061	3.916	1	0.048	***
PER3	-0.094	PER11	-0.015	4.628	1	0.031	***
PER4	-0.045	PER12	0.004	1.620	1	0.203	
PER5	-0.066	PER13	-0.028	0.609	1	0.435	
PER6	-0.015	PER14	0.043	1.486	1	0.223	

PER7	-0.116	PER15	-0.108	0.022	1	0.883
PER8	-0.072	PER16	-0.046	0.211	1	0.646
HKG1	-0.411	HKG9	-0.351	0.941	1	0.332
HKG2	-0.433	HKG10	-0.403	0.133	1	0.715
HKG3	-0.399	HKG11	-0.324	1.845	1	0.174
HKG4	-0.377	HKG12	-0.369	0.013	1	0.911
HKG5	-0.328	HKG13	-0.326	0.001	1	0.980
HKG6	-0.330	HKG14	-0.350	0.038	1	0.844
HKG7	-0.343	HKG15	-0.382	0.718	1	0.397
HKG8	-0.329	HKG16	-0.355	0.254	1	0.614

Note. ABS=absence from school, NOR= Norway, PER= Peru, HKG= Hong Kong, SS= statistically significant.

Appendix 9 a

Percentages (%) of Females Classified as Disadvantaged or Academically Resilient in Norway, Peru, and Hong Kong

Student	Student	Norway			Peru			Hong Kong		
Background	**Outcome**	**ALL**	**DIS**	**RES**	**ALL**	**DIS**	**RES**	**ALL**	**DIS**	**RES**
Fixed ESCS	Fixed	49.60	52.81	57.67	49.63	49.85	36.22†	49.92	49.20	49.35
	Relative			62.26			42.11†			44.69†
Fixed WEALTH	Fixed		50.36	54.55		50.74	41.91†		49.15	49.93
	Relative			46.88			45.49†			47.34
Fixed BOOKS	Fixed		43.35*#	41.26		48.96	36.93†		45.05*#	46.05
	Relative			32.41†			42.25†			40.09
Fixed EMOSUPS	Fixed		45.73*#	45.76		49.36	37.53†		47.23*	48.26
	Relative			41.36			42.78†			46.11
Relative ESCS	Fixed		51.48	52.49		51.02	28.44†		48.57	49.19
	Relative			51.63			39.34†			45.06
Relative WEALTH	Fixed		52.91*	53.61		51.02	34.96†		48.72	49.65

	Relative			50.45			43.12†			46.55
Relative BOOKS	Fixed		43.55*#	40.65		49.57%	32.57†		45.22*	46.42
	Relative			31.21†			39.53†			40.39
Relative EMOSUPS	Fixed		45.66*#	45.77		49.69%	37.94†		46.73*	48.18
	Relative			41.31			44.10†			46.80

Appendix 9 b

Percentages (%) of Students With a Different Language Classified as Disadvantaged or Academically Resilient in Norway, Peru, and Hong Kong

Student Background	**Student Outcome**		**Norway**		**Peru**			**Hong Kong**		
		ALL	**DIS**	**RES**	**ALL**	**DIS**	**RES**	**ALL**	**DIS**	**RES**
Fixed ESCS	Fixed	8.72	25.77*	19.02	7.17	9.60*	1.59†	3.51	4.03	1.42
	Relative			9.43†			2.07†			0.28
Fixed WEALTH	Fixed		22.63*	16.88		8.31*#	1.17†		4.02	1.62
	Relative			9.38			1.62†			0.63

Fixed BOOKS	Fixed	18.59*#	13.23†	8.68*	0.95†	5.07*	2.28†
	Relative		11.03†		1.77†		1.11†
Fixed EMOSUPS	Fixed	10.60*#	6.96†	8.76*	1.86†	3.33	1.89†
	Relative		5.56†		2.19†		1.25†
Relative ESCS	Fixed	14.98*	11.00†	14.56*	3.67†	4.41	1.70†
	Relative		8.97†		4.26†		0.43†
Relative WEALTH	Fixed	15.88*	11.14†	15.32*	4.07†	5.12*	1.93†
	Relative		9.22†		4.59†		0.93†
Relative BOOKS	Fixed	15.18*	10.62†	9.85*#	0.00†	5.35*	2.30†
	Relative		7.01†		1.40†		1.22†
Relative EMOSUPS	Fixed	10.66*#	7.00†	9.19*#	2.01†	4.04	1.82†
	Relative		5.72†		2.28†		1.07†

Note. ALL= percentage of females/students who spoke another language at home in all, DIS= percentages of females/ students who spoke another language at home in the group of students classified as disadvantaged, RES=percentage of females/ students who spoke another language at home in the group of students classified as academically resilient, *= percentage of females/students who spoke another language at home in disadvantaged students is statistically significantly different from percentage of females/students who spoke another language in all, † = percentage of females/students who spoke another language in resilient students is statistically significantly different from percentage of females/students who spoke another language in disadvantaged students, #= the proportion of disadvantaged students under operationalizations of WEALTH, BOOKS, or EMOSUPS was significantly different from the proportion of disadvantaged students under the operationalization of ESCS using the same threshold.

Appendix 10 a

Mean Performance of Females Classified as Disadvantaged or Academically Resilient in Norway, Peru, and Hong Kong

		Norway				**Peru**				**Hong Kong**			
		TOP 1/3		**RES**		**TOP 1/3**		**RES**		**TOP 1/3**		**RES**	
Back	**Out**	**Male**	**Fem**	**Male**	**Fem**	**Male**	**Fem**	**Male**	**Fem**	**Male**	**Fem**	**Male**	**Fem**
Fixed ESCS	Fixed	560.96	**550.77***	526.26	530.34	513.64	**508.44***	503.91	498.96	561.00	**555.12***	554.83	**546.56***
	Relative	607.13	**600.90***	582.69	580.77	484.55	**477.89***	472.62	**465.24***	610.53	**604.88***	606.26	601.81
Fixed WEALTH	Fixed	560.96	**550.77***	544.36	533.15	513.64	**508.44***	507.99	503.38	561.00	**555.12***	557.23	**551.15***
	Relative	607.13	**600.90***	590.88	594.74	484.55	**477.89***	478.51	**472.00***	610.53	**604.88***	607.89	**602.84***
Fixed BOOKS	Fixed	560.96	**550.77***	531.57	**522.41***	513.64	**508.44***	505.54	**498.43***	561.00	**555.12***	548.21	**539.66***
	Relative	607.13	**600.90***	582.96	582.69	484.55	**477.89***	475.59	**466.34***	610.53	**604.88***	603.97	599.17
Fixed EMOSUPS	Fixed	560.96	**550.77***	555.60	**544.55***	513.64	**508.44***	513.55	507.50	561.00	**555.12***	558.00	552.20
	Relative	607.13	**600.90***	607.55	600.01	484.55	**477.89***	485.07	**474.49***	610.53	**604.88***	608.77	**602.01***
Relative ESCS	Fixed	560.96	**550.77***	544.52	**537.09***	513.64	**508.44***	499.20	498.47	561.00	**555.12***	551.85	**545.83***
	Relative	607.13	**600.90***	598.23	**587.93***	484.55	**477.89***	466.12	**458.75***	610.53	**604.88***	605.44	601.91

Relative WEALTH	Fixed	560.96	**550.77***	562.91	**551.63***	513.64	**508.44***	495.99	498.38	561.00	**555.12***	557.99	**548.91***
	Relative	607.13	**600.90***	612.25	**603.70***	484.55	**477.89***	465.00	460.64	610.53	**604.88***	609.60	**601.49***
Relative BOOKS	Fixed	560.96	**550.77***	543.32	**527.23***	513.64	**508.44***	504.98	**493.49***	561.00	**555.12***	549.52	**540.48***
	Relative	607.13	**600.90***	595.73	**585.70***	484.55	**477.89***	471.77	**459.88***	610.53	**604.88***	603.80	599.16
Relative EMOSUPS	Fixed	560.96	**550.77***	554.84	**544.11***	513.64	**508.44***	514.66	**506.37***	561.00	**555.12***	555.27	551.60
	Relative	607.13	**600.90***	607.09	600.00	484.55	**477.89***	486.61	**473.65***	610.53	**604.88***	605.19	600.19

Appendix 10 b

Mean Performance of Students Who Spoke a Different Language at Home Classified as Disadvantaged or Academically Resilient in Norway, Peru, and Hong Kong

		Norway				**Peru**				**Hong Kong**			
		TOP 1/3		**RES**		**TOP 1/3**		**RES**		**TOP 1/3**		**RES**	
Background	**Out**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**
	Fixed	557.35	**541.08***	533.52	**509.54***	511.44	510.79	502.07	505.12	558.88	**528.14***	551.28	**512.01***

Fixed ESCS	Relative	605.05	594.45	583.31	564.11	481.76	474.06	469.63	463.76	608.09	**586.34***	604.33	**573.54***
Fixed WEALTH	Fixed	557.35	**541.08***	543.57	512.03	511.44	510.79	505.99	511.34	558.88	**528.14***	554.75	**519.63***
	Relative	605.05	594.45	591.67	602.55	481.76	474.06	475.71	465.74	608.09	**586.34***	605.68	**573.95***
Fixed BOOKS	Fixed	557.35	**541.08***	529.07	519.73	511.44	510.79	502.86	507.96	558.88	**528.14***	544.90	**517.17***
	Relative	605.05	594.45	584.22	571.97	481.76	474.06	471.92	458.39	608.09	**586.34***	602.34	**571.74***
Fixed EMOSUPS	Fixed	557.35	**541.08***	551.33	540.65	511.44	510.79	511.24	513.48	558.88	**528.14***	555.68	**531.28***
	Relative	605.05	594.45	604.00	611.76	481.76	474.06	480.71	473.04	608.09	**586.34***	605.86	**590.25***
Relative ESCS	Fixed	557.35	**541.08***	542.12	**529.09***	511.44	510.79	499.00	498.69	558.88	**528.14***	549.44	**514.18***
	Relative	605.05	594.45	593.82	583.74	481.76	474.06	463.53	456.18	608.09	**586.34***	603.94	**573.54***
Relative WEALTH	Fixed	557.35	**541.08***	558.68	**542.40***	511.44	510.79	496.32	508.65	558.88	**528.14***	554.28	**514.84***
	Relative	605.05	594.45	608.87	598.72	481.76	474.06	463.14	462.88	608.09	**586.34***	606.11	**575.03***
Relative BOOKS	Fixed	557.35	**541.08***	538.65	**521.37***	511.44	510.79	501.24	0.00	558.88	**528.14***	545.92	**519.78***
	Relative	605.05	594.45	593.34	582.81	481.76	474.06	467.53	**434.94***	608.09	**586.34***	602.25	**571.74***
Relative EMOSUPS	Fixed	557.35	**541.08***	550.66	540.76	511.44	510.79	511.43	515.73	558.88	**528.14***	554.07	**524.08***
	Relative	605.05	594.45	603.70	611.76	481.76	474.06	481.80	474.80	608.09	**586.34***	603.36	583.93

Note. Out=outcome, TOP 1/3= top 1/3 student, RES= academically resilient students, L1=students who spoke test language at home, L2=students who spoke another language at home, *= mean performance of females/students who spoke another language at home is statistically significantly different from mean performance of males/ student who spoke test language at home in the same operationalization.

Appendix 11 a

Proportion of Females and Males in Top 1/3 Students in Science Achievement

	Norway					Peru					Hong Kong				
	TOP 1/3	Male		Female		TOP 1/3	Male		Female		TOP 1/3	Male		Female	
Out	n	n	%	n	%	n	n	%	n	%	n	n	%	n	%
Fixed	3467	1735	50.04%	1732	49.96%	1305	756	57.93%	549	42.07%*	4189	2052	48.99%	2137	51.01%
Relative	1780	949	53.31%	831	46.69%*	2387	1315	55.09%	1072	44.91%*	1882	962	51.12%	920	48.88%*

Appendix 11 b

Proportions of Students Who Spoke Another Language at Home in Top 1/3 Students in Science Achievement

	Norway						Peru						Hong Kong					
	top1/3	L2		All	L2		top1/3	L2		All	L2		top1/3	L2		All	L2	
Out	n	n	%	n	n	%	n	n	%	n	n	%	n	n	%	n	n	%
Fixed	3396	207	6.10%†	5456	476	8.72%	1303	21	1.64%†	6971	500	7.17%	4125	79	1.92%†	5359	188	3.51%

Relative	1752	87	4.97%†	5456	476	8.72%	2377	45	1.93%†	6971	500	7.17%	1864	20	1.07%†	5359	188	3.51%

Note. Out= students outcome, top 1/3= top 1/3 students in science achievement, All= all students in a country, L2= students who spoke another language at home, *= the proportion of female was significantly different from the proportion of male within the same operationalization, †=the proportion of students who spoke another language at home in top 1/3 was significantly different from the proportion of students who spoke another language at home in all.

Article 3

Ye, W., Olsen, R. V., & Blömeke, S. (2023). More money does not necessarily help: Relations of education expenditure, school characteristics, and academic resilience across 36 education systems (under review in the journal of Large-scale Assessments in Education)

3

Abstract

Teacher quality, teaching quality, school resources, and school climate are commonly identified as protective factors in the academic resilience literature. Variables reflecting these four concepts were applied in a latent profile analysis across thirty-six education systems participating in the Trends in International Mathematics and Science Study 2019. The best-fitting model suggested four different latent profiles of protective factors. A three-step BCH method with an auxiliary regression model was adopted to investigate the influence of education expenditure on academic resilience across the profiles. Education expenditure promoted academic resilience in a profile characterized by low mathematics resources and another profile with low teaching quality and school climate. Education expenditure had no significant influence in the remaining two profiles characterized by very low and high levels of classroom and school protective factors, respectively. Moreover, countries were classified into six cultural groups representing education systems sharing similarities in language, history, or geography. Within each group, there was a certain degree of consistency in the distribution of profiles. Conclusions are drawn for strategies to promote academic resilience.

Keywords. Academic resilience, Teacher quality, Teaching quality, School resources, School climate, Education expenditure

1. Introduction

Educational inequalities are a concern in many countries across the globe. With the increasing availability of data from international large-scale assessments (ILSAs), a growing number of studies have examined how to promote “academic resilience”, a term that refers to succeeding “against the odds”. Academic resilience describes students’ capacity to perform well despite having a disadvantaged background (OECD, 2011). The critical question is what characterizes malleable features of the school contexts that are positively related to such academic resilience (Agasisti et al., 2018) and whether it would be possible to increase these features by increasing education expenditure.

The growing utilization of ILSAs data offers researchers valuable insights into the protective factors derived from students’ individual characteristics, family backgrounds, and learning environments. Moreover, it presents a significant opportunity to examine the role of these protective factors in fostering academic resilience across countries (i.e., Cheung, 2017). However, when exploring academic resilience across countries, three primary concerns emerge.

First, researchers examining protective factors across nations usually explore their overall influences by either analyzing one pooled data set or multiple single-country data sets. The former approach usually ignores country-specific characteristics (i.e., Agasisti et al., 2018). The latter often employs a single model to investigate the influence of protective factors within individual countries, leading to results that are challenging to generalize (i.e., Erberer et al., 2015).

Second, the operationalization of academic resilience is problematic in many studies, as shown by Authors. (2021). International comparison studies often utilize a fixed performance threshold to operationalize academic resilience, which refers to a single score used to denote exceptional educational achievement across all participating nations. This approach fails to consider the wide variations in average academic achievement across countries and may result in an inadequate understanding of what it means to be a high achiever in countries at the lower end of the achievement scale.

Third, multilevel modeling was usually adopted to explore the relationships between academic resilience and protective factors. Researchers frequently explored numerous protective factors and solely reported the relationship between academic resilience and protective factors without delving into interactions (i.e., Vicente et al., 2021). Consequently, the intricate relationships and nonlinear associations among these protective factors were often overlooked. To address this issue, some researchers (i.e., Koirikivi et al., 2021) utilized latent class analysis to examine multiple protective factors, aiming to identify distinct classes of resilience resources. These identified classes were subsequently examined with external variables using a three-step method. However, this method is not without limitations, either. Specifically, when the external variable (i.e., outcome) is included in the final stage, the latent class variable may experience substantial shifts in membership, thereby rendering the results invalid (Asparouhov & Muthén, 2014b).

These problems call for a more suitable conceptualization of academic resilience and a more comprehensive evaluation of protective factors across countries. The current study employs data from 8^{th} graders, teachers, and principals within 36 education systems participating in the Trends in International Mathematics and Science Study (TIMSS) 2019. As compared to other ILSAs, the sampling design of TIMSS allows for directly linking student and teacher data. Moreover, the grade 8 population, as compared to the grade 4 population in TIMSS, has much more robust information on students' home backgrounds. The study aims to identify resilient students and to analyze how profiles of protective factors vary across nations with state-of-the-art methods avoiding the problematic issues identified above. Moreover, the variability across countries is used to explore how these profiles of protective factors reflect educational expenditure.

1.1 Academic Resilience

Resilience refers to positive adaptation despite adversity (Luthar, 2006). Depending on the measurement method, studies utilize certain adverse characteristics to define risk; hence positive adaptations refer to outcomes better than expected. Protective factors, which facilitate resilience, are a fundamental research topic in the field (Tudor & Spray, 2017).

When resilience is explored in education, positive adaptations usually focus on students' academic performance (OECD, 2011), while risks are defined in various ways (Martin & Marsh, 2008). Studies on academic resilience went through two periods and demonstrated different patterns: before and after using ILSAs data. In the first period, researchers treated problematic relationships with parents or discrimination as student risks (Wayman, 2002) while emphasizing individual protective factors such as persistence (Martin & Marsh, 2008). Studies focused on aspects associated with students and their families, such as students' attitudes toward school and family academic support (Wayman, 2002).

With the development of ILSAs, standardized information about students' knowledge and skills became available across countries. Accordingly, studies investigating *academic* resilience in an international comparative context have become frequent since the 2010s. ILSAs also came with composite measures of students' socio-economic home background, such as PISA's economic, social and cultural status (ESCS). Thus many studies adopted a low level of SES to define risk. Moreover, with a growing awareness that schools can compensate for risks such as a disadvantaged home background, emphasis was placed on malleable institutional factors such as school resources, school climate, and teaching styles and strategies (Agasisti et al., 2018). Considering the impact of education financial policy on these amenable school inputs, some researchers further examined the influence of education expenditure on academic resilience (Agasisti & Longobardi, 2012).

Data used in this period usually included one country or more, and naturally, researchers started wondering whether the protective factors' impact varies across countries. With the help of ILSAs data, international comparisons became more accessible. However, due to variations in economic development, cultures, and educational policies, researchers face significant challenges when defining and operationalizing academic resilience as a unitary concept across countries. In studies that utilize ILSAs data, a relative threshold is typically employed to define risk (i.e., bottom 1/3 of SES within-country), and a fixed cut-off on the scale of educational achievement is used to determine positive adaptation (i.e., a score of 475 defined to be the lower bound of the achievement level labeled as "intermediate" in the TIMSS studies). This approach may lead to overestimating the proportion of resilient students in high-achieving countries and underestimating it in low-achieving countries, which may not accurately reflect the quality of support provided to their disadvantaged students. This study, therefore, adopted a relative threshold to determine high educational performance, as Authors. (2021) suggested, to better reflect the pool of academically resilient students in each country.

1.2 Factors Promoting Academic Resilience

1.2.1 School Characteristics

Research on academic resilience is deeply rooted in the fields of psychology and sociology (Aburn, Gott, & Hoare, 2016). Hence, the protective factors most frequently studied reflected either within-person traits/states or characteristics reflecting the level of society. Accordingly, research on resilience did not until recently to a large extent reflect the potential of the teacher, and other resources proximal to the instructional context, as potential protective factors facilitating resilience. For example, extra-curricular activities were found to promote resilience for adolescents experiencing behavioral or mental health difficulties (Sun, 2007).

The state of research changed with Borman and Overman (2004), who thoroughly examined protective factors related to core school characteristics in the United States. Borman and Overman emphasized school inputs, such as school resources (free-lunch legibility, availability of instructional resources, class size), teacher quality (years of experience), curriculum and instructional quality (clear goals, monitoring student progress), and school climate (safe and orderly environment). Their research revealed that characteristics of a supportive school community, including a safe and orderly environment, positive teacher-student relationships, and support for family involvement, were the most influential factors in promoting academic resilience.

The emphasis on malleable school factors was further underscored when ILSAs data was applied to study resilience. Using Italian data from PISA 2009, Agasisti and Longobardi (2012) focused on school-level characteristics. The study found that school factors associated with teachers were generally significant in predicting resilience, including the availability of teaching resources, the proportion of qualified teachers, the teacher-student ratio, and teacher shortage.

Erberer and colleagues (2015) employed TIMSS 2011 data from twenty-eight countries to investigate the relationships between academic resilience and school characteristics. They studied teachers' beliefs that students can do well in mathematics, the percentage of disadvantaged students, schools' emphasis on academic success, safety and discipline, and shortages in educational resources on instruction. The associations between school factors and resilience were found to vary across education systems, with the most robust and consistent predictor being the beliefs held by teachers.

Utilizing combined data from PISA 2012 and 2015, Agasisti et al. (2018) examined the associations between academic resilience and various school-level factors. The factors included school learning climate (disciplinary climate, percentage of students skipping school days, extra-curricular activities), school resources (computer-student ratio, class size, average school SES), and school leadership. Only the computer-student ratio was not significantly associated with academic resilience.

García-Crespo and colleagues (2021) used data from the Progress in International Reading Literacy Study (PIRLS) 2016 to investigate student resilience in reading, emphasizing teachers' influences. These included teachers' formal education level and specialization, the school's emphasis on academic success, a safe and orderly school environment, teacher-student interaction, teachers' job satisfaction, classroom instructional limitations due to student attributes, reading strategies and techniques, homework tracking, and selection of reading materials. Findings showed that effective classroom management, a safe and orderly school environment, and teaching methods were the top predictors of academic resilience. In a follow-up study, García-Crespo and colleagues (2022) used TIMSS 2019 data and investigated teaching-related variables in mathematics and science. Schools' emphasis on academic success and a safe and orderly climate predicted academic resilience across domains.

To sum up, four core characteristics of educational quality in schools were widely discussed in the literature using data from ILSAs: teacher quality, teaching quality, school climate, and school resources. The indicators of these were often positively related to academic resilience. However, it is essential to note that using ILSAs data to investigate academic resilience has resulted in varying research focus areas due to distinct sample designs in different ILSAs. For example, the PISA data does not associate students with their teachers because students were randomly selected from sample schools (OECD, 2020). Thus, studies using PISA data to explore teacher factors tend to examine them at the school level, i.e., the proportion of qualified teachers in a school (Jin et al., 2022). This has limited the depth of knowledge regarding the influence of teachers and teaching quality on academic resilience.

1.2.2 Education Expenditure

As a significant determinant of school inputs, education expenditure reflects the necessary financial resources for schools to establish a conducive learning environment, enhance teacher quality, and provide adequate school resources. Although education expenditure was found to have a limited direct average influence on student achievement (Hanushek & Woessmann, 2017), it may help *disadvantaged* students perform better.

Agasisti and Longobardi (2014a) employed PISA 2009 OECD data to investigate school factors related to teachers (i.e., teacher shortage) in conjunction with education expenditure and institutional characteristics. They found that education expenditure, the number of teaching hours per year, teachers' average salary after 15 years of experience, and the age at which students were first grouped by ability were positively related to resilience.

Following these outcomes, Agasisti and Longobardi (2014b) explored the association between the percentage of resilient students and two types of education expenditures: education spending as a fraction of a) government expenditure and b) GDP. Their study revealed that the former was associated with higher academic resilience in OECD countries.

However, the latter was found to function in a compensatory manner, with a slightly negative association observed for richer countries but a positive one for poorer countries.

Agasisti and colleagues (2017) extended their inquiry into the relationship between education spending as a part of government expenditure and the presence of academic resilience in the OECD context by scrutinizing the data from five PISA cycles (2000 to 2012). Their findings revealed that education expenditure as a percentage of government spending might assist disadvantaged students, but the magnitude and direction of this association could be contingent on a country's level of economic development. Specifically, it was observed as beneficial in poorer nations but unfavorable in richer ones.

Vicente and colleagues (2021) employed PISA data from 2003 to 2018 to investigate the influence of individual factors (i.e., self-confidence), school factors (i.e., school SES, class size), and country-level factors (i.e., education expenditure per student, the ratio of teacher salary and GDP) on academic resilience. Their research revealed that, in the case of poorer countries, education expenditure per student is a significant predictor for academic resilience, whereas, for richer nations, teacher salary can contribute to enhanced academic performance among disadvantaged students.

To sum up, based on studies of PISA data, academic resilience seems to correlate positively with education expenditure, particularly in poorer countries. Furthermore, these studies suggest that the efficacy of expenditure depends on the allocation and utilization of funds. The interrelationships among education expenditure, core teaching and teacher characteristics, and academic resilience remain inadequately understood due to the constraints imposed by PISA data. Specifically, the inability to establish a direct linkage between students and their teachers limits studies utilizing PISA data to scrutinize resilience solely at the school level. As a result, the complexity and nuances of teacher-related factors have not been fully explored in the literature.

1.3 General versus Country-Specific Influences of Protective Factors

As mentioned in the introduction, studies exploring protective factors enhancing academic resilience across countries can be distinguished by two trends: (1) examination of general influences using pooled data and (2) comparative analyses across a few selected countries. The former approach produced more significant results across studies due to a larger sample size but was less context-specific, while findings from the latter were often challenging to generalize. Additionally, the results from both approaches could vary considerably based on the levels and covariates included in the analysis.

1.3.1 Examination of General Influences Using Pooled Data

Using pooled data from OECD PISA 2009, Agasisiti and colleagues (2014a) established the significance of education expenditure to academic resilience, which was confirmed and expanded upon in their subsequent research (Agasisiti et al., 2014b). Their research on school-level factors, such as the influence of the computer-student ratio on resilience, was consistent across studies based on pooled data.

However, associations between academic resilience and protective factors differed when various covariates and levels were considered in the analysis. The significant association between education expenditure and school-level factors diminished when educational systems' characteristics were considered in the model (Agasisti et al., 2012). Similar observations were made regarding the hierarchical level of analysis utilized in the model. The significant connections between academic resilience and school-related variables, such as extra-curriculum activities, were no longer evident upon incorporating country-level factors, such as education expenditure, into the model (Agasisiti et al., 2014a).

Including multiple countries' data from ILSAs in one pooled dataset may contribute to the complexity of interpreting results, as the substantial differences in cultural and educational contexts may significantly impact the assessment outcome. For instance, the proportion of teachers who have completed bachelor's or postgraduate degrees is highly variable, ranging from 100% in Canada to 1% in Saudi Arabia (Mullis et al., 2020). Therefore, pooled data analysis may fail to account for these contextual differences accurately, and the significance of the findings may be obscured or misinterpreted.

1.3.2 Comparative Analyses Across Countries

Considering the heterogeneity among nations, some researchers investigated academic resilience country by country or in a few selected countries only. Erberer et al. (2015) examined protective factors in twenty-eight education systems participating in TIMSS 2011. Their findings indicated that factors at the individual level, such as students' academic aspirations, demonstrated greater consistency than school-level factors, such as the school's emphasis on academic success. García-Crespo et al. (2021) conducted a series of studies comparing protective factors across countries and found inconsistent results. For example, school discipline was significantly related to academic resilience in only eight out of twenty-three countries.

Given the inconsistent outcomes of studies exploring protective factors across countries, some researchers have opted to limit their investigations to countries with shared characteristics such as culture, language, geographic location, or levels of economic development. Cheung et al. (2014) studied the influence of individual factors, namely, enjoyment of reading, diversity of reading materials, and metacognitive awareness of reading strategies, on reading resilience across four East Asia economies. The study yielded consistent results across countries for all three variables. Subsequently, Cheung (2017) reported consistent results concerning mathematics learning variables across five education systems in East Asia. Nevertheless, Sandoval-Hernández and colleagues (2016) reported inconsistent results when analyzing individual and school-level characteristics in five East Asian economies. For example, students' valuing of mathematics was a predictor of academic resilience in three out of five education systems, whereas school emphasis on academic success was significant in only one country.

Meanwhile, some researchers also compared protective factors between countries with notable cultural and economic differences. Ni and colleagues (2018) compared elementary students' resilience in China and the United States. They found considerable country differences in individual resilience-promoting characteristics like self-control. Gabrielli and colleagues (2022) investigated protective factors between Southern European and North-western countries. They found that school-level factors like extra-curricular activities were not significantly related to academic resilience in the former group, but the opposite was found in the latter.

Özcan (2022) extended the research on academic resilience by comparing the influences of protective factors between individualist and collectivist cultures. Their study revealed significant disparities in the influences of individual characteristics, such as the self-concept of reading. However, negligible variations were observed concerning school-level resilience-promoting factors such as disciplinary climate.

1.4 Methodological Challenges

Compared to studies employing pooled data, research utilizing country-specific data faces more difficulties concerning the methodologies applied in the analyses. ILSAs data like TIMSS typically involve a sample size of approximately 4,000 students per country, with a comparatively lower representation of disadvantaged students (Mullis & Martin, 2017). Smaller sample sizes decrease not only statistical power but also the flexibility of the effect size, which can result in the study being unable to detect a significant difference or correlation between variables, even if one is present (Anderson, Kelley, & Maxwell, 2017).

Given the hierarchical structure of schools, researchers commonly construct a baseline logistic regression model that includes individual-level variables and subsequently introduces classroom or school-level factors (Agasisti et al., 2018). However, the relationships among covariates may differ across countries, resulting in issues with model convergence. Consequently, multilevel modeling investigations often fail to explore interactions and primarily report the connections between protective factors and academic resilience. As a result, the interdependence among covariates across different hierarchical levels and its influence on the relationship between protective factors and academic resilience are frequently overlooked in these studies.

To address these methodological challenges, Putwain and colleagues (2013) adopted cluster analysis and examined factors related to academic resilience across groups of students. Subsequently, Collie and colleagues (2017) expanded on this method by utilizing analysis of variance to establish associations between clusters of resilience and students' motivation.

The progress in statistical techniques has enabled researchers to establish the relationship between latent class membership and external variables. One approach that has gained widespread use in recent years is the three-step approach proposed by Asparouhov and Muthén (2014a). This approach involves estimating the latent class measurement model and subsequently examining the association between the latent class variable and the auxiliary variables. Boutin-Martinez et al. (2019) and Koirikivi et al. (2021) employed latent class analysis to identify distinct profiles of resilience-related factors, taking into account covariates and using the three-step method to explore auxiliary variables.

However, empirical studies have suggested that the three-step method may not fully address the issue of shifting classes[1], particularly when the entropy[2] value is low and there is a significant disparity in the variances of the distal outcome across classes (Bakk and Vermunt, 2014). To address this problem, Asparouhov and Muthén (2014b) proposed the BCH[3] method, which avoids shifts in latent class in the final stage to which the three-step method is susceptible. Despite the potential advantages of the BCH method, no resilience-related studies have yet, to the best of our knowledge, utilized this method.

[1] The introduction of external variables into the regression analysis may result in alterations to the probability of individuals being assigned to specific latent classes.

[2] Entropy is a statistical fit index for model-based classification accuracy, with higher values indicating more precise assignment of individuals to latent profiles. Generally, a value close to 1 is ideal and above .8 is acceptable.

[3] Named after Bolck, Croon and Hagenaars who developed this method.

1.5 The Present Study

In this study, country-specific characteristics are taken into account as relative thresholds for risks and positive adaptations are employed to define academic resilience. Taking advantage of the TIMSS design, in which students are nested in classrooms, we examine teacher and school characteristics previously documented as promising protective factors.

Using latent profile analysis (LPA), this study investigates protective factors patterns by grouping the observed classroom and school characteristics into distinct profiles of resilience resources, to help improve our understanding of how such resources may work differently across cultures. Adopting the BCH method, we further examined the extent to which profiles are associated with academic resilience and education expenditure via auxiliary regression models. Specifically, we address the following research questions:

1. How many distinct profiles of resilience resources, characterized by teacher quality, teaching quality, school climate, and school resources, can be identified in the sample?
2. Do the profiles of resilience resources exhibit identifiable cultural patterns across diverse nations?
3. To what extent do the identified latent profiles predict academic resilience?
4. To what extent does the association between education expenditure as a percentage of Gross Domestic Product (GDP) and academic resilience vary across profiles?

2. Method

2.1 Sample and Procedure

To better understand teachers' impact on academic resilience, we utilized TIMSS 2019 data, in which one or more intact classes were selected from randomly sampled schools via a two-stage stratified cluster sampling design (Mullis & Martin, 2017). The unprocessed data comprises 1.33 teachers per school[4], with the majority of schools having only one teacher. In order to optimize a parsimonious model, one single teacher was randomly selected from schools represented by more than one teacher. T-tests were subsequently employed to compare teacher-related variables between the original and the modified dataset with only one teacher per school. Except for teachers' educational level and specialization in Finland, and teachers' educational specialization in Chinese Taipei, no other statistically significant

[4] The TIMSS data encompass one or more classrooms per school, with some instances involving multiple math teachers per school. However, the majority of the unprocessed data in our study consist of one classroom per school and one math teacher per classroom.

differences were detected in the remaining items that were tested (details see Tables 1-5 in Supplementary Materials).

Following the appropriate treatment of 11 variables[5] employed for latent profile identification and the random selection of one teacher per school, this study applied relative thresholds to define disadvantaged students within each education system. The final sample used for the present study thus comprises eighth-grade disadvantaged students (N = 54,748) and their mathematics teachers (n = 6,798) and principals (k = 6,798) from 36 educational systems participating in TIMSS 2019 Mathematics. Analyses were thereby conducted on this sub-sample of disadvantaged students and their teachers and principals.

2.2 Measures[6]

2.2.1 Academic Resilience

This study adopted relative thresholds for risk and positive adaptation. Low SES (bottom 1/3 within-country) was used to define risk, and high mathematics performance (top 1/3 within-country) was used for positive adaptation. The Home Educational Resource (HER) scale of TIMSS 2019, based on students' answers about home possessions and parents' highest level of education, was used to indicate students' SES. A higher score refers to a higher level of SES (see Table 6 in Supplementary Materials).

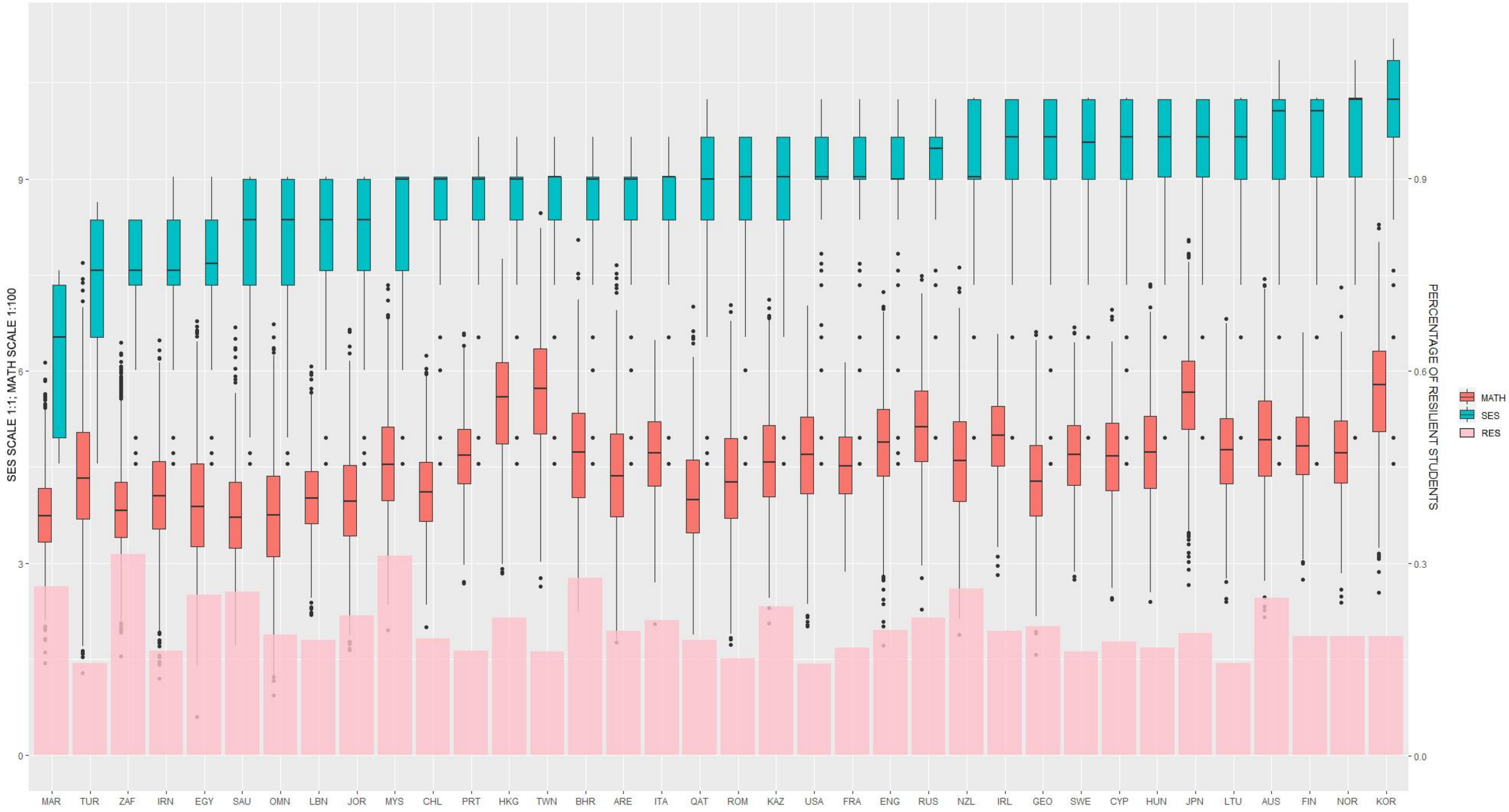
The publicly available data sets from TIMSS represent students' achievement by five plausible values (PVs) where students' achievement scores are conditioned on all available background data (Mullis & Martin, 2017). This study used all five plausible values in mathematics, following Rubin's (1987) rules. After identifying disadvantaged students as specified above, those with high performance were defined as resilient (1); otherwise, as non-resilient (0). Consequently, the five plausible values were converted into five binary numbers that each assumed a value of either 1 or 0. These five binary numbers were used as dependent outcomes in the last step of the analysis. In M*plus* software, this is achieved by specifying "TYPE = IMPUTATION" in the data command.

With relative performance thresholds of 561.88 and 414.08, the United States and South Africa exhibited the lowest and highest occurrence of resilient students, respectively (see Figure 1).

[5] Students' ratings on teaching quality (classroom management and instructional quality) were aggregated to the classroom level, and teachers' ratings on school's emphasis on academic success and safe and orderly climate were aggregated to the school level.

[6] Most of the variables used in this paper have composite scale scores in TIMSS, which the technical report confirms are comparable. We conducted measurement invariance tests on the remaining scores.

Figure 1 *SES, Mathematics Performance, and Percentage of Resilience in Disadvantaged Students Across 36 Education Systems*



Note. SES = Social-economic status, RES = Percentage of Resilient students, MATH = Mathematics performance/100.

Table 1 *Summary of Classroom and School Protective Factors (11 items)*

	Label	Description	Raters	Type
Teacher Quality	BTBG04*	Level of Formal Education Complete	Teachers	7-point Likert scaled score
	BTDMMME*	Majored in Mathematics and Mathematics Education	Teachers	5-point Likert scaled score
Teaching Quality	CAgSCALE	Cognitive Activation Activities (based on items BTBG12A-12G)	Teachers	Composite scale score based on 7 items
	BSBGICM*	Instructional Clarity in Mathematics Lessons (aggregated to the classroom level)	Students	Composite scale score derived from TIMSS
	BSBGDML*	Disorderly Behavior during Math Lessons/Classroom Management (aggregated to the classroom level)	Students	Composite scale score derived from TIMSS
School Resources	SCHSES	Students' HER aggregated to the school level	Students	Mean
	SCHOTL	taught curriculum topics/all curriculum topics	Teachers	Fraction
	BCBGMRS*	Instruction Affected by Mathematics Resources Shortage	Principals	Composite scale score derived from TIMSS
School Climate	BCBGDAS*	School Discipline Problems	Principals	Composite scale score derived from TIMSS
	BTBGEAS*	School Emphasis on Academic Success	Teachers	Composite scale score derived from TIMSS
	BTBGSOS*	Safe and Orderly Schools	Teachers	Composite scale score derived from TIMSS

Note. SCHSES = school socio-economic status, SCHOTL= school opportunity to learn, * = generated by TIMSS, HER = Home Educational Resource, except for BTBG04 and BTDMMME, all other nine variables are continuous; this study tested measurement invariance for the self-calculated variable, CAgSCALE (cognitive activation), but not for the others.

2.2.2 Teacher Quality

This study assessed teacher quality by their highest level of education attained and classification of which subject or field was their major field of study. Teachers were asked to report their highest level of formal education completed on a scale ranging from 1 (did not complete upper secondary education) to 7 (Doctor or equivalent level), with a higher score indicating a higher level of educational attainment. Regarding the majors of the teachers, we used the variable, “Teachers Majored in Mathematics and Mathematics Education,” provided by the TIMSS data set. The scale of this variable ranges from 1 (major in mathematics and mathematics education) to 5 (no formal education in mathematics beyond upper secondary), which was reversed in our study such that a higher score denotes a stronger formal background in mathematics and education (see Table 1).

2.2.3 Teaching Quality

This study scrutinized three teaching quality dimensions: cognitive activation, instructional clarity, and classroom management (Blömeke, Olsen, & Suhl, 2016). Cognitive activation was evaluated through teachers’ responses to “How often do you do the following in teaching this class.” Seven items were rated on a four-point Likert scale, with examples including “related the lesson to students’ daily lives”. Given the involvement of numerous educational systems in the present study, we utilized relaxed fit indices to assess measurement invariance for cognitive activation (Nagengast & Marsh, 2014). Specifically, we relied on changes in the Comparative Fit Index (CFI) below .01 and changes in the Root Mean Square Error of Approximation (RMSEA) at or below .015. With ΔCFI = 0.00 and ΔRMSEA = 0.014, scalar invariance was established. Subsequently, we constructed a scale (α = .780) based on these items, with a higher score reflecting greater levels of cognitive activation.

For instructional clarity and classroom management, this study employed two TIMSS scales based on student responses, which were subsequently aggregated to the classroom level for appropriate modeling purposes. The scale of Instructional Clarity in Mathematics Lessons comprised seven items, including statements such as “My teacher is easy to understand,” with a higher score denoting higher levels of instructional clarity. The scale of Disorderly Behavior During Mathematics Lessons, consisting of six items, such as “there is disruptive noise,” was used as a measure of classroom management, with a higher score indicating fewer incidents of disorderly behavior. Composite scores for these two scales generated by TIMSS were used in the analysis.

2.2.4 School Resources

Three aspects of school resources were investigated: mathematics resources as reported by principals, school SES as a proxy of overall school resources, and opportunity to learn (OTL) mathematics as reported by teachers for the target class.

TIMSS's composite scale score of the Instruction Affected by Mathematics Resources Shortage, which was derived from five items such as "Library resources relevant to mathematics instruction", was used to measure mathematics resources. A higher score indicates that mathematics instruction is less impacted by resource shortage. School SES was created on students' HER scores, aggregated to the school level. School OTL, which measures the extent to which students have access to high-quality curriculum, was based on teachers' responses to 22 topics related to Numbers, Algebra, Geometry, and Data and Probability. If the teacher indicated that the specific topic, for instance, "Simple linear equations" was "Mostly taught before this year" or "Mostly taught this year", the topic was regarded as taught. The number of topics taught divided by the total number of 22 topics was used as a measure of OTL.

2.2.5 School Climate

This study examined three aspects of school climate: discipline, schools' emphasis on academic success (SEAS), and a safe and orderly environment.

The composite scale score on School Discipline Problems was derived from principals' answers on 11 items such as "Absenteeism", with a higher score indicating a lower incidence of disciplinary problems. The other two composite scale scores—SEAS and Safe and Orderly School—were based on teachers' responses. A higher score indicates a higher level of academic emphasis or a safer and more orderly climate. The SEAS scale was derived from 14 items such as "Teachers' expectation for student achievement". The Safe and Orderly School scale was based on eight items such as "I feel safe at this school." All three composite scale scores were developed by TIMSS and designed to provide comparable measures across countries.

2.2.6 Education Expenditure

This study adopted the Government Expenditure on Education (% of GDP) data from the World Bank, calculated by dividing total government expenditure for all levels of education by GDP. In the analyses, averages of government expenditure on education from 2016 to 2018, three years before TIMSS 2019, were employed, ranging from 2.26% to 7.87%. By averaging the data over three years, we can reduce the impact of year-to-year fluctuations and obtain a more stable and reliable measure for education expenditure. Saudi Arabia and the three Nordic countries, Norway, Sweden, and Finland, ranked highest in education expenditure.

2.3 Statistical Analyses

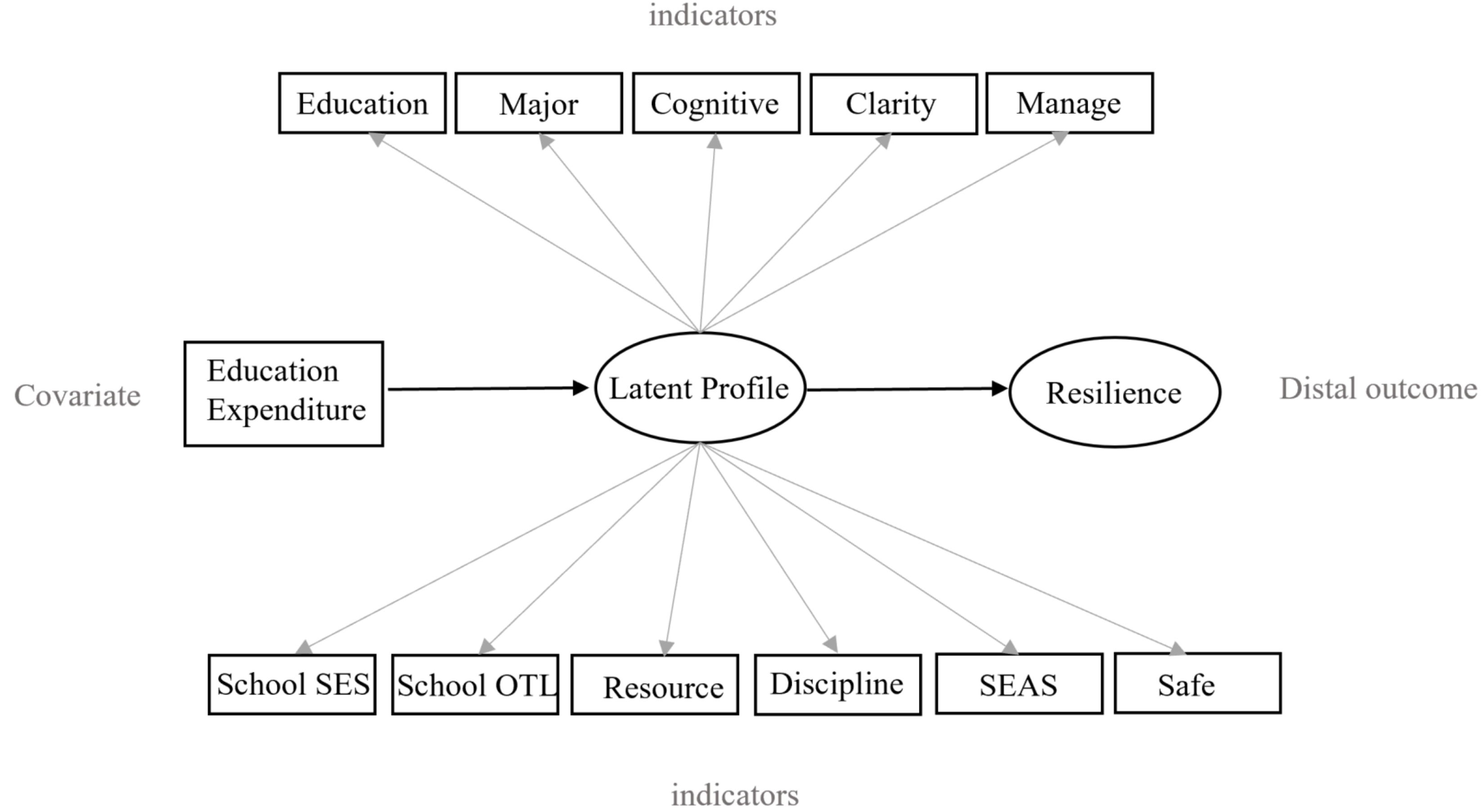
Data was prepared using the R Software version 4.0.2 (R Core Team, 2019). Analyses were conducted using M*plus* version 8.8 (Muthén & Muthén, 2017), and missing data were handled with the full information maximum likelihood. Maximum likelihood estimation with robust standard errors (MLR) was used for latent profile analyses. TIMSS 2019 data are available from the International Association for the Evaluation of Education Achievement (IEA) database at https://timss2019.org/international-database/. Education expenditure data can be downloaded from the World Bank website at https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS.

Asparouhov and Muthén (2014a) proposed a three-step method to establish a connection between latent classes and external variables while considering measurement errors. This method involves: 1) identifying latent classes; 2) classifying memberships (calculating the average classification error for each identified class); and 3) linking the identified classes with external variables. To address issues related to shifting class membership, Asparouhov and Muthén (2014b) further developed the BCH method. This method closely resembles the three-step approach but differs in the second step, where the classification error is calculated for each individual.

The present study employs latent profile analysis (LPA) to identify multiple unobserved latent homogenous profiles. To investigate the influence of education expenditure on academic resilience across these profiles, we adopted a BCH method, which allows for linkage between the profiles and covariates, as well as with distal outcomes (Asparouhov & Muthén, 2014b).

Figure 2 provides the conceptual model employed in the current study. The 11 school and classroom characteristics used as indicators in Step 1 were employed to identify the optimal latent profile model. In Step 2, individual classification errors were computed, and the inverse logits of these error rates were used as BCH weights in the next step (Nylund-Gibson, Grimm & Masyn, 2019). In Step 3, education expenditure and academic resilience were incorporated into the model as the covariate and the distal outcome, respectively.

Figure 2 *Conceptual Model of the Latent Profile Analysis and Auxiliary Regression*



Note. SES = socio-economic status, OTL= opportunity to learn, SEAS = School’s emphasis on academic success.

Based on preliminary analyses and theoretical considerations, covariances among these 11 items were also included in the latent profile analysis (LPA). Four models were assessed to determine the suitable LPA model. Model 1 assumed equal variances across profiles and fixed covariances to zero. Model 2 entailed equal variances and covariances across profiles. Model 3 involved freely estimated variances and equal covariances across profiles. Lastly, Model 4 allowed both freely estimated variances and covariances across profiles.

Comparisons across models of the following fit indices were used to decide the appropriate number of latent profiles: the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the sample-size adjusted BIC (SABIC), the Lo-Mendell-Rubin likelihood ratio test (LMR), and the bootstrap likelihood ratio test (BLRT). Besides, we also considered parsimony and theoretical meaningfulness (Nylund, Asparouhov, & Muthén, 2007).

After identifying the latent profiles of resilience resources, two auxiliary logistic regressions were performed using BCH weights derived from Step 2. The first regression examined the relationship between academic resilience and identified profiles, while the second regression incorporated the covariate variable, namely education expenditure, to assess the influence of education expenditure on academic resilience across profiles. Given that the BCH setting is not applicable in M*plus* for multilevel design, we specified "TYPE = COMPLEX" and "CLUSTER = SCHOOLID" in the command to account for the hierarchical structure of students nested within schools.

3. Results

3.1 Research Question 1: Profiles of Resilience Resources

3.1.1 Identifying Profiles of Resilience Resources

Eleven classroom and school protective factors were used as latent profile indicators. In the following, we refer to these as resilience resources. Four distinct model configurations, each characterized by differing covariances and variances, were employed to determine the optimal LPA model. Model 1 exhibited entropy values of approximately 0.6, indicating a relatively low level of classification quality and suggesting inadequate fit of the model to the observed data (Wang et al., 2017). Models 3 and 4 had convergence problems when the profile numbers increased. This issue is common with less restrictive LPA models with many free parameters that may lead to unstable solutions (Bauer, 2022). Therefore, Model 2, which assumed equal variances and covariances across profiles, was chosen as the preferred model.

As the number of profiles increased, the AIC, BIC, and SABIC values decreased according to the LMR and BLRT tests (see Table 2). The decline in statistical significance was not evident in the five-profile model, indicating that the four-profile solution provides the optimal fit for the data, despite the two-profile model having the highest entropy value. Differences in entropy were limited, though. The smallest profiles in all models were smaller than 5%, ranging from .809% to 1.765%. Still, for the four-profile model, the smallest profile had 104 schools, higher than the recommended 50 units (Weller, Bowen, & Faubert, 2020). We also found that the smallest group in the four-profile model made reasonably conceptual sense. Therefore, the model with four profiles was chosen as both supported by model fit statistics and conceptual considerations.

Table 2 *Model Fit Indices for Latent Profile Analyses*

Profiles	**AIC**	**BIC**	**SABIC**	**Entropy**	**LMR p-value**	**BLMR p-value**	**Smallest profile%**
1	274637.592	275163.07	274918.382				
2	272423.568	273030.938	272748.118	.988	.000	.000	1.765
3	270433.591	271122.854	270801.900	.923	.000	.000	1.589
4	267693.930	268465.085	268105.998	.939	.000	.000	1.530
5	267549.272	268402.319	268005.100	.939	.993	1.000	.809

Note. N = 6,798, AIC = Akaike's Information Criterion, BIC = Bayesian Information Criterion, SABIC = Sample-Adjusted BIC, LMR = Lo-Mendell-Rubin, BLRT = Bootstrap Likelihood Ratio Test.

3.1.2 Description of the Four Profiles of Resilience Resources

Figure 3 presents the latent profiles of the four groups across the 11 resilience resources, with the y-axis representing the standardized value for these resources. As commonly experienced in such latent profile analysis, one group represents those with low values across the indicators, while another group represents a profile with overall high average values across the indicators. In our case, profile 1, stands out as substantially different from the other three profiles with extremely low scores on all resilience resources. In contrast, profile 4 represents a group of students with overall high values across all 11 resilience resources. The two other profiles also had overall relatively high values for many resources but with some substantially lower values for some indicators.

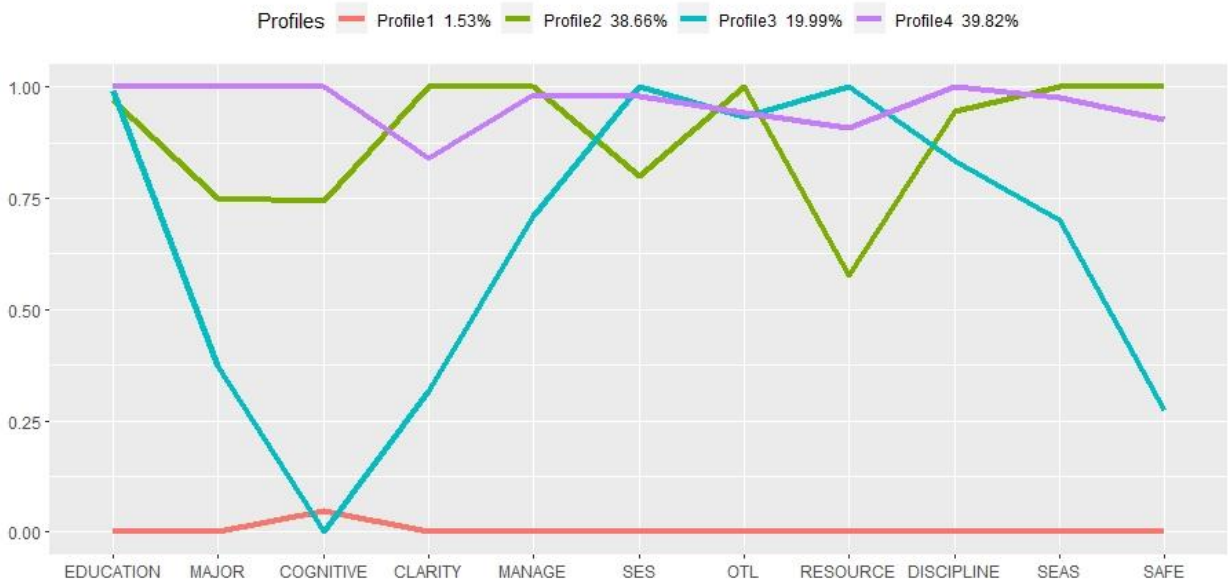
In the following more detailed descriptions for each of the profiles are provided. Profile 1 had the lowest group size (1.53%), characterized by the lowest levels of teacher quality, teaching quality (instructional clarity and classroom management), school resources, and school climate. Hence, this profile was labeled “Vulnerable”. It should also be noted that in addition to being a rather small group, this profile is also mostly defined by one country (Morocco, see Figure 4). Hence, this profile appears partly as being an outlier in the solution.

Profile 2 was the second-largest group (38.66%). Schools in this profile demonstrated high teaching quality, as evidenced by high instructional clarity and good classroom management. They also had a positive school climate, with the highest ratings for safe and orderly climate and schools’ emphasis on academic success. However, it ranked relatively low in mathematics resources but provided nevertheless a lot of OTL. This profile was named “Effective Teaching and Positive Climate.”

Profile 3, with a group size of 19.99%, was characterized by a high level of school resources, including the highest school SES and mathematics resources. However, it had a low level of teaching quality, characterized by the lowest level of cognitive activation and relatively lower levels of instructional clarity and classroom management. Fewer teachers in Profile 3 had a major in mathematics and education. Furthermore, the school climate received relatively low ratings across all indicators, encompassing disciplinary measures, the school’s emphasis on academic success, as well as safety and orderly climate. The profile was denoted as “Resource-Heavy, Quality-Light.”

Profile 4 had the largest group size (39.82%). Teachers in this profile exhibited the highest level of teacher quality on both education level and subject-specific major, and a high level of teaching quality on all other indicators, cognitive activation, instructional clarity, and classroom management. Schools in this profile also had a good school climate and ranked high on school resources, including school SES, OTL, and mathematics resources. For brevity, this profile was named “Good Schools.”

Figure 3 *Plots of Four Latent Profiles*



Note. SES = socio-economic status, OTL= opportunity to learn, SEAS = School's emphasis on academic success.

3.2 Research Question 2: Cultural Patterns in Resilience Resources Profiles Across Countries[7]

Schools characterized by a vulnerable profile (Profile 1) were found in 23.53% of the schools in Morocco, 4.90% in Sweden, and 4.35% in Jordan. The prevalence of the profile in the other countries, excluding the 21 out of 36 economies where Profile 1 was absent, ranged from 0.41% to 3.66% (see Figure 4A).

The frequency of Profile 2, characterized by effective teaching and a favorable school environment, varied across countries, with Jordan exhibiting the highest proportion of schools (72.17%), followed by Romania (70.11%) and Saudi Arabia (66.33%). In contrast, the United States (12.06%), Hungary (8.90%), and Portugal (8.97%) demonstrated the lowest proportions of schools classified as Profile 2.

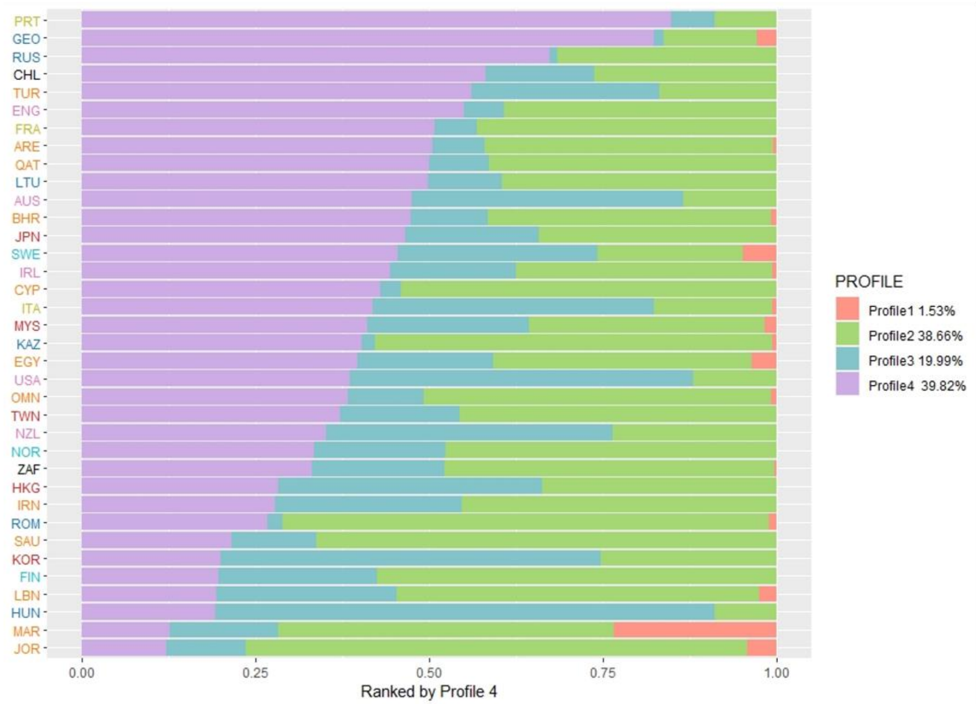
The profile labeled "Resource-Heavy, Quality-Light" (Profile 3) exhibited a lower average prevalence across nations (19.99%). Among the economies studied, the highest proportion was observed in Hungary (71.92%), followed by South Korea (54.82%) and the United States (49.42%). In contrast, Romania (2.17%), Kazakhstan (1.89%), Georgia (1.42%), and Russia (1.02%) had the lowest prevalence of Profile 3.

[7] Classification for each school was based on its most likelihood of membership.

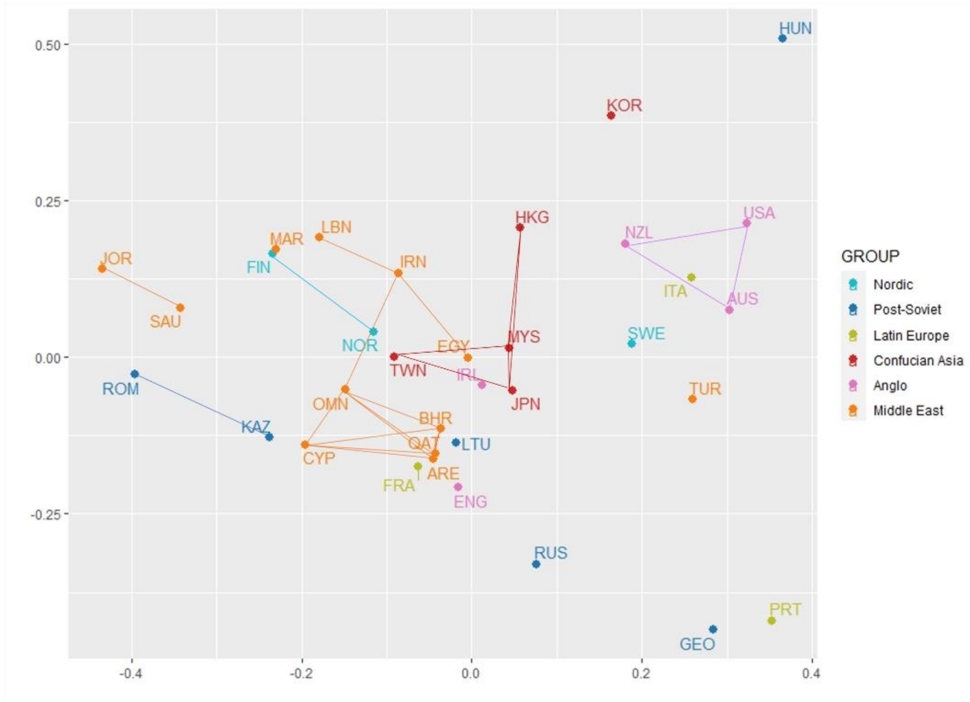
Figure 1

The Presence of Four Latent Profiles in 36 Education Systems

(A) The Presence of Four Latent Profiles in 36 Education Systems



(B) Multidimensional Scaling Plot Based on the Presence of Four Latent Profiles



Note. For education system codes, see Table 3; In Plot (B), lines between economies indicate no statistically significant difference.

Table 3 *Cultural Groups*

Groups	Code	Education Systems	%GDP	Groups	Code	Education Systems	%GDP
Middle East (12)	ARE	United Arab Emirates	*NA*	**Anglo (5)**	NZL	New Zealand	6.240
	SAU	Saudi Arabia	7.871		ENG	United Kingdom	5.309
	CYP	Cyprus	5.932		AUS	Australia	5.182
	OMN	Oman	5.524		USA	United States	4.940
	MAR	Morocco	5.234		IRL	Ireland	3.548
	TUR	Turkey	4.294	**Latin Europe (3)**	FRA	France	5.429
	EGY	Egypt	4.050		PRT	Portugal	4.845
	IRN	Iran, Islamic Rep. of	3.704		ITA	Italy	4.039
	JOR	Jordan	3.222	**Nordic (3)**	NOR	Norway	7.862
	QAT	Qatar	2.967		SWE	Sweden	7.609
	BHR	Bahrain	2.520		FIN	Finland	6.497
	LBN	Lebanon	2.260	**Post-Soviet (6)**	HUN	Hungary	4.617
Confucian Asia (5)	TWN	Chinese Taipei	*NA*		RUS	Russian Federation	4.376
	MYS	Malaysia	4.635		LTU	Lithuania	3.903
	KOR	Korea, Rep.of	4.373		GEO	Georgia	3.563
	HKG	Hong Kong SAR	3.310		ROM	Romania	3.140

JPN	Japan	3.118	KAZ	Kazakhstan	2.782

Note. Chile (CHL, %GDP = 5.40) and South Africa (ZAF, %GDP = 5.56) were not grouped, *NA* = not available.

The “Good schools” (Profile 4) were most prevalent in Portugal at 84.83%, Georgia at 82.27%, and Russia at 67.35%. Conversely, the lowest presence was observed in Hungary (19.18%), Morocco (12.61%), and Jordan (11.30%).

As depicted in Figure 4A, the share of these four profiles displayed substantial variation across nations. Nonetheless, when cultural differences were taken into account, certain patterns emerged. This study categorized 34 education systems into six groups based on geographic and cultural considerations (see Table 3). We adopted a broader definition for the Middle East group, including countries in the Middle East and North Africa (Akkari, 2004). The Confucian Asia group included East and Southeast Asian education systems belonging to the Confucian cultural sphere (Huang & Chang, 2017). All countries in the Anglo group are developed nations, predominantly English-speaking, and once British colonies (Ashkanasy, Trevor-Roberts, & Earnshaw, 2002). The Latin Europe group included three Western or Southern European countries in which Romance languages are predominant. Three countries in the Nordic group share an egalitarian idea that the education system should provide access and opportunities for all (Frønes et al., 2020). The post-Soviet group includes six countries that used to be dominated by or part of the Soviet Union. Since Chile and South Africa were culturally and geographically different from these clusters, they were not included in any of the groups.

The two-dimensional Figure 4B was constructed using the four profiles present in each economy. The *ggplot2* package in R (Wickham et al., 2016) was utilized to generate a multidimensional scaling plot, a statistical technique employed to present complex, high-dimensional data in a lower-dimensional space, while preserving the original distances or similarities as much as possible. In Figure 4B, closer proximity between economies indicated greater similarity in their four profiles. The present study utilized Fisher’s exact test to evaluate the similarities within cultural groups. The lines in Figure 4B indicated no statistically significant difference between the economies compared. We discuss our results, therefore, within cultural groups.

Seven Middle East countries were found to have Profile 1 schools, which were observed in only 15 economies. Except for Turkey, Middle East countries demonstrated a relatively high prevalence of Profile 2, roughly 40% or greater. In contrast, Profile 3 was less prevalent, ranging from 27.04% to 7.52%. Profile 4 formed two distinct clusters characterized by higher and lower percentages. The latter cluster comprised Jordan, Morocco, Lebanon, and Saudi Arabia, most classified as low-SES countries. Fisher’s exact test results indicate no significant differences among Oman, Bahrain, United Arab Emirates, Qatar, and Cyprus (p = .577). In contrast, Turkey and Morocco exhibited significant differences from the other Middle East countries.

Profile 1 was absent in Confucian Asia, except for Malaysia. Profile 2 had a similar prevalence (around 34%) in Hong Kong, Malaysia, and Japan, with Chinese Taipei having a higher prevalence and South Korea having a lower prevalence. Profile 3 was more prevalent in South Korea (54.82%) and Hong Kong (37.90%), while the other three had lower proportions (< 23.21%). Profile 4 was less common in South Korea (19.88%) and Hong Kong (28.23%) but more common in the other three systems (> 37.06%). Fisher's test results showed no statistically significant differences among Malaysia, Japan, and Chinese Taipei (p = .381) or between Malaysia, Japan, and Hong Kong (p = .056). However, South Korea was significantly different from the other Confucian Asia economies.

Except for Ireland, Profile 1 was not found in Anglo countries. The prevalence of Profile 2 was relatively lower in Anglo countries than in other groups, with the United Kingdom (39.34%) and Ireland (36.91%) having moderate representation. Profile 3 had a lower presence in Ireland (18.12%) and the United Kingdom (5.74%) but higher shares in the other three (>38.93%). Profile 4 had moderate to high representations in Anglo countries, ranging from 35.11% to 54.92%. In this group, Fisher's test results found no significant differences among New Zealand, the United States, and Australia (p = .104).

For the Latin Europe group, no cultural pattern was identified. Profile 1 was found in Italy but not the other two. Portugal and Italy had low shares of Profile 2, but France had a relatively high proportion (43.15%). Profile 3 had a higher presence in Italy (40.52%) but lower in the other two (< 6.21%). All three countries had a moderate to high share of Profile 4, ranging from 41.83% to 84.83%.

Among Nordic countries, Sweden showed significant differences compared to Norway and Finland. Regarding resilience resource profiles, Sweden exhibited a unique pattern with the presence of Profile 1, a relatively low proportion of Profile 2 (20.98%), and a relatively high ratio of Profile 4 (45.45%), respectively, as compared to the other two countries. Fisher's test results also supported this finding. Norway and Finland were not significantly different (p = .082).

Profile 1 was present in Georgia, Kazakhstan, and Romania. Profile 2 had a higher share in Kazakhstan (57.23%) and Romania (70.11%), a moderate presence in Russia and Lithuania (around 30%), and a low share in Georgia and Hungary (< 13.48%). Except for Hungary

(71.92%), Profile 3 had the lowest presence in Post-Soviet countries, ranging from 1.02% to 2.17%. Profile 4 had high shares in Georgia (82.27%) and Russia (67.35%), moderate shares in Lithuania and Kazakhstan (around 40%), and relatively low shares in Romania and Hungary (about 20%). Within this group, the distribution of the profiles for Romania and Kazakhstan was not statistically different (p = .071).

3.3 Research Question 3: Relationship Between Academic Resilience and Identified Profiles

Following the completion of Steps 1 and 2, this study initially conducted an auxiliary logistic regression analysis, regressing the distal outcome of academic resilience on the four identified profiles. The purpose of this analysis was to investigate how these profiles predict academic resilience.

In the regression analysis examining the association between latent profiles and the binary outcome of academic resilience, the odds ratios were obtained from the M*plus* output. These odds ratios were then utilized to calculate the corresponding probabilities of being resilient across the different latent profiles (see Table 4), enabling a more accessible interpretation of the results. The probability estimates for students being resilient in Profiles 1, 2, and 3 were calculated as 0.221, 0.213, and 0.206, respectively. In contrast, the probability of students being resilient in Profile 4, referred to as “Good Schools,” was estimated to be 0.281. We further tested probabilities across profiles via model constraint. The results revealed that the probability of being resilient in Profile 4 was significantly higher than the probabilities observed in the other three profiles (p = .000).

Table 4 *Relationship Between Four Latent Profiles and Academic Resilience*

		95% C.I.			
	Odds Ratio	**Lower2.5%**	**Upper2.5%**	***p*-value**	**Probability**
Profile 1	0.961	0.818	1.128	0.000	0.221
Profile 2	1.064	1.041	1.088	0.000	0.213
Profile 3	1.061	1.004	1.121	0.000	0.206
Profile 4	1.035	0.958	1.118	0.000	0.281

Note. C.I. = Confidential Interval.

3.4 Research Question 4: Relationship Between Academic Resilience and Education Expenditure Across Profiles

Subsequently, we introduced the covariate, namely education expenditure, into the auxiliary logistic regression, investigating its association with academic resilience across profiles. Positive associations were observed between academic resilience and education expenditure in profiles 2, 3, and 4, with statistically significant relationships found only in profiles 2 (p = .000) and 3 (p = .037). In contrast, Profile 1 exhibited a negative but non-significant association (p = .627) between education expenditure and academic resilience.

Profile 1, denoted as "Vulnerable," displayed an odds ratio of 0.961 (95% CI [0.818, 1.128]). Profile 2, labeled as "Effective Teaching and Positive Climate," demonstrated an odds ratio of 1.064 (95% CI [1.041, 1.088]), signifying a 6.4% higher likelihood of resilience with a one-unit increase in education expenditure. Profile 3, termed "Resource-Heavy, Quality-Light," exhibited an odds ratio of 1.061 (95% CI [1.004, 1.121]), indicating a 6.1% higher likelihood of resilience with a one-unit increase in education expenditure. Profile 4, referred to as "Good Schools," yielded an odds ratio of 1.035 (95% CI [0.958, 1.118]).

Table 5 *Relationship Between Education Expenditure and Academic Resilience*

					95% C.I.	
	Estimate	**S.E.**	***p*-value**	**Odds Ratio**	**Lower2.5%**	**Upper2.5%**
Profile 1	-0.04	0.096	0.627	0.961	0.818	1.128
Profile 2	0.062	0.016	0.000	1.064	1.041	1.088
Profile 3	0.059	0.037	0.037	1.061	1.004	1.121
Profile 4	0.034	0.058	0.389	1.035	0.958	1.118

Note. S.E. = Standard Error, C.I. = Confidential Interval.

4. Discussion

Prior research has identified classroom and school characteristics as predictors of academic resilience. However, certain technical and practical constraints, such as the design of ILSAs and limited access to SES data, have led to a lack of exploration regarding the influential role of these characteristics, including teacher and teaching quality, school resources, and school climate, in predicting academic resilience. Moreover, it has been an open question whether it is possible to identify cultural patterns in such protective factors.

Scholars have dedicated attention to investigating the impact of education expenditure, particularly in countries with limited resources. However, research on the effectiveness of education expenditure in relation to school inputs has been sparse, and there remains a dearth of knowledge regarding the allocation and utilization of funds. Additionally, concerns have been raised regarding the validity of operationalizations of academic resilience, which may impede the generalization of findings across studies.

This study was designed to examine whether profiles of resilience resources exist across different countries, whether cultural patterns can be identified, and whether education expenditure promotes academic resilience within these identified profiles.

4.1 Profiles of Resilience Resources

This study employed latent profile analysis (LPA), a less commonly utilized approach in the field, to investigate protective factors associated with academic resilience from four key perspectives: teacher quality, teaching quality, school resources, and school climate. By incorporating the covariances among 11 items representing these perspectives, the study accounted for country-specific characteristics in identifying latent profiles of resilience resources. While direct relationships among these protective factors were not established through LPA, several patterns can be observed between them.

One interesting characteristic of the four profiles is that where the teacher quality (particularly when their speciliaztions are more related to mathematics) is high, the three indicators of teaching quality also are high. This is consistent with findings from previous studies that teacher quality and teaching quality are associated (Blömeke et al., 2016).

Another notable characteristic observed in the four profiles is the co-occurrence of low teaching quality, encompassing cognitive activation, instructional clarity, and classroom management, with a correspondingly low school climate. Specifically, indicators reflecting the school’s emphasis on academic success, as well as the climate of order and safety, exhibit lower levels when teaching quality is diminished. This phenomenon can be partially elucidated by the findings of Gore et al. (2022), who discovered that differences in teaching quality are less a reflection of teacher capabilities than of challenging circumstances. A better school climate may mitigate disturbances during class, thereby enabling teachers to focus on improving teaching interactions rather than classroom management.

Prior research has established a positive association between school resources and teaching quality (Hill, Blazar, & Lynch, 2015). Schools with better resources were also known to attract and retain qualified teachers. However, the current study reveals a contrasting pattern wherein profiles characterized by greater school resources, encompassing school socio-economic status (SES), school opportunities to learn (OTL), and mathematics resources, do not consistently exhibit higher levels of teacher quality and teaching quality. This finding may be attributed to educational equity policies that seek to redistribute accomplished teachers to socioeconomically disadvantaged schools.

As previously indicated, profiles characterized by higher levels of school climate exhibit correspondingly higher levels of teacher quality and teaching quality. It is noteworthy that the most pronounced disparities are observed in the indicator measuring a safe and orderly climate. This observation can be attributed to the notion that a safe and orderly climate serves as a supportive environment for teachers, fostering their emotional well-being, social interactions, and academic endeavors (García-Crespo et al., 2021).

4.2 Cultural Patterns in Resilience Resources Profiles Across Countries

Previous research has assumed that countries sharing similar backgrounds tend to exhibit similarities in academic resilience and have consequently explored protective factors across various countries, including East Asian countries (Cheung et al., 2014) and Southern European countries (Gabrielli et al., 2022). However, there is a dearth of empirical studies investigating the underlying reasons for this assumption. To address this gap, the current study examined the presence of four identified profiles of resilience resources across six distinct cultural groups.

Consistent with prior investigations, several of the examined cultural groups demonstrated certain cultural similarities. Specifically, five out of twelve Middle East countries, four out of five Confucian Asian economies, three out of five Anglo countries, two out of three Nordic countries, and two out of six Post-Soviet countries exhibited no significant differences in their respective profiles of classroom and school protective factors.

However, noteworthy differences were also observed within these cultural groups, particularly among Middle East and Post-Soviet countries. Specifically, the presence of Profile 4, labeled as “Good Schools,” exhibited distinct proportions in two clusters in Middle East countries, one characterized by higher SES and the other characterized by lower SES. These differences may be partially attributed to variations in economic development, suggesting that the observed disparities in the prevalence of Profile 4 could be influenced by economic factors.

4.3 Academic Resilience and Profiles of Resilience Resources

Consistent with a substantial body of prior research (Agasisti et al., 2018; Garcia-Crespo et al., 2022), this study revealed a similar pattern whereby students in profiles characterized by higher levels of resilience resources, including teacher quality, teaching quality, school resources, and school climate, exhibited a correspondingly higher likelihood of academic resilience. In particular, students in Profile 4, distinguished by the highest ratings across nearly all indicators pertaining to the four perspectives of resilience resources, demonstrate the highest probability of exhibiting resilience.

However, an interesting finding emerged where students belonging to Profile 1, denoted by the lowest level of resilience resources, displayed the second highest probability of exhibiting resilience. This phenomenon can be attributed to the operationalization of academic resilience employed in this study, which adopted relative thresholds for performance. Notably, Profile 1 predominantly consisted of students from Morocco, a country characterized by the lowest mean SES among the 36 educational systems in our sample (see Figure 1). Despite this, Morocco demonstrated a relatively high prevalence of resilient students, ranking within the top five among the 36 education systems studied.

4.4 Academic Resilience and Education Expenditure

Consistent with prior investigations (Agasisti, 2017), our analysis revealed significant positive associations between education expenditure and academic resilience for two out of four profiles. Within Profile 2, characterized by low mathematics resources, and Profile 3, characterized by low teaching quality and school climate, a positive and statistically significant association was observed between education expenditure and academic resilience. This finding suggested that promoting academic resilience through these specific factors is plausible.

However, the influence of education expenditure on academic resilience was found to be statistically non-significant in Profile 4, characterized as "Good Schools." This observation can be attributed to the already high quality of the schools within this profile, suggesting that additional education expenditure did not result in further improvements in academic resilience. An alternative explanation could be that the education expenditure allocated within this profile was directed toward educational aspects not captured by ILSAs, such as initiatives aimed at promoting student well-being or investments in music and arts education.

In Profile 1 ("Vulnerable"), the relationship between education expenditure and academic resilience exhibited a negative but statistically non-significant association. Although not reaching statistical significance, this negative association provides insights into potential disparities in the distribution of education resources within education systems including Profile 1 schools. Further empirical investigations are warranted to substantiate this assumption and deepen our understanding of resource allocation dynamics in these contexts.

5. Limitations

With an emphasis on profiling resilience resources among disadvantaged students, this study exclusively scrutinized the teacher and school-related information pertaining to this subgroup. It is important to note that this limited scope may not comprehensively represent overall school inputs, particularly in countries characterized by significant income distribution disparities.

Moreover, disadvantaged students were identified based on those who remained in school. Yet, it should be acknowledged that some of the most vulnerable students may already have dropped out of school and are not included in the sampling frame. The application of relative thresholds for risk and positive adaptation yields relatively higher proportions of resilient students in low-SES countries, which could be partially explained by lower coverage of the target population in those countries.

To explore classroom factors associated with students' academic resilience, we used the data from the TIMSS study, in which students are nested in a classroom. Given parsimonious model considerations, classroom and school factors were incorporated both at the school level in this study. Nonetheless, it is important to note that teacher quality in Finland and Taipei exhibited notable disparities between the raw data and the data utilized in this study. Consequently, careful scrutiny is necessary when interpreting the results of these two education systems, particularly concerning teacher education levels and majors.

Furthermore, it should be noted that using the BCH method in this study has methodological limitations that preclude the application of multilevel modeling, including multilevel LPA. While schools were specified as the cluster to address the nesting structure of students and schools, treating education expenditure as a school-level variable does not capture the variances that may exist across schools.

Although this paper examined the government expenditure on education (%GDP), no information was known about how funds were utilized. It is plausible that some countries may allocate funds to expand the scope of education, including areas such as music and arts, which are not typically evaluated by ILSAs. Besides, we adopted the mean of 2016 to 2018 to measure education spending, which may not accurately reflect the change across time. Moreover, it took time for protective factors to influence the outcome. Future studies are encouraged to explore protective factors longitudinally.

6. Conclusion

This study found that profiles with teachers more likely to have a major in mathematics and education tend to have favorable teaching qualities. Furthermore, a higher level of teaching quality was observed in profiles with more favorable school climates. However, school resources were not necessarily associated with teacher quality, teaching quality, and school climate in these identified profiles. The presence of these identified four latent profiles varied across cultures. Several cultural patterns were found, which confirms the assumption that similar countries may be more alike in academic resilience. Profiles with higher level of resilience resources tended to predict academic resilience. The varied associations between education expenditure and academic resilience across profiles underscore the significance of contextual factors in supporting disadvantaged students. For instance, providing a dependable avenue for accessing resources might constitute a crucial prerequisite for assisting disadvantaged students in vulnerable schools. Alternatively, in good schools, factors such as effective teaching might be more influential in supporting disadvantaged students. These findings underscore the importance of tailoring interventions and policies to specific contexts and demand greater attention to the complexity of promoting academic resilience in schools.

References

Aburn, G., Gott, M., & Hoare, K. (2016). What is resilience? An integrative review of the empirical literature. *Journal of advanced nursing*, *72*(5), 980-1000. https://doi.org/10.1111/jan.12888

Agasisti, T., & Longobardi, S. (2014a). Inequality in education: Can Italian disadvantaged students close the gap? *Journal of Behavioral and Experimental Economics*, *52*, 8-20. http://dx.doi.org/10.1016/j.socec.2014.05.002

Agasisti, T., & Longobardi, S. (2017). Equality of educational opportunities, schools' characteristics and resilient students: An empirical study of EU-15 countries using OECD-PISA 2009 data. *Social Indicators Research*, *134*(3), 917-953. https://doi.org/10.1007/s11205-016-1464-5

Agasisti, T., Longobardi, S., & Regoli, A. (2014b). Does public spending improve educational resilience? A longitudinal analysis of OECD-PISA data. *Working Papers 3, Società Italiana di Economia Pubblica.*

Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). Academic resilience: What schools and countries do to help advantaged students succeed in PISA. *OECD Education Working Papers*, *167*. https://doi.org/10.1787/e22490ac-en

Akkari, A. (2004). Education in the Middle East and North Africa: The current situation and future challenges. *International Education Journal*, *5*(2), 144-153.

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science, 28*(11), 1547-1562. https://doi.org/10.1177/0956797617723724

Asparouhov, T., & Muthén, B. (2014a). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural equation modeling: A Multidisciplinary Journal*, *21*(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Asparouhov, T., & Muthén, B. (2014b). Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes, 21*(2), 1-22. Retrieved from: https://www.statmodel.com/examples/webnotes/webnote21.pdf

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology, 43*(1), 272-311. https://doi.org/10.1177/0081175012470644

Bauer, J. (2022). A primer to latent profile and latent class analysis. In *Methods for Researching Professional Learning and Development: Challenges, Applications and Empirical Illustrations* (pp. 243-268). Cham: Springer International Publishing.

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher Quality, Instructional Quality and Student Outcomes*. Cham: Springer International Publishing, *2*, 21-50.

Borman, G. D., & Overman, L. T. (2004). Academic resilience in mathematics among poor and minority students. *The Elementary School Journal*, *104*(3), 177-195.

Boutin-Martinez, A., Mireles-Rios, R., Nylund-Gibson, K., & Simon, O. (2019). Exploring resilience in Latina/o academic outcomes: A latent class approach. *Journal of Education for Students Placed at Risk, 24*(2), 174-191. http://dx.doi.org/10.1080/10824669.2019.1594817

Cheung, K. C. (2017). The effects of resilience in learning variables on mathematical literacy performance: A study of learning characteristics of the academic resilient and advantaged low achievers in Shanghai, Singapore, Hong Kong, Taiwan and Korea. *Educational Psychology, 37*(8), 965-982. https://doi.org/10.1080/01443410.2016.1194372

Collie, R. J., Ginns, P., Martin, A. J., & Papworth, B. (2017). Academic buoyancy mediates academic anxiety’s effects on learning strategies: an investigation of English- and Chinese-speaking Australian students. *Educational Psychology*, *37*(8), 947-964. https://doi.org/10.1080/01443410.2017.1291910

Drent, M., Meelissen, M. R., & van der Kleij, F. M. (2013). The contribution of TIMSS to the link between school and classroom factors and student achievement. *Journal of curriculum studies*, *45*(2), 198-224. https://doi.org/10.1080/00220272.2012.727872

Frønes, T. S., Pettersen, A., Radišić, J., & Buchholtz, N. (2020). Equity, equality and diversity in the Nordic model of education—Contributions from large-scale studies. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-030-61648-9_1

Gabrielli, G., Longobardi, S., & Strozza, S. (2022). The academic resilience of native and immigrant-origin students in selected European countries. *Journal of Ethnic and Migration Studies, 48*(10), 2347-2368. https://doi.org/10.1080/1369183X.2021.1935657

García-Crespo, F. J., Fernández Alonso, R., & Muñiz, J. (2021). Academic resilience in European countries: The role of teachers, families, and student profiles. *Plos one, 16*(7), e0253409. https://doi.org/10.1371/journal.pone.0253409

García-Crespo, F. J., Fernández Alonso, R., & Muñiz, J. (2022). Academic resilience in mathematics and science: Europe TIMSS-2019 Data. *Psicothema, 34*(2), 217-225. http://dx.doi.org/10.7334/psicothema2021.486

Gore, J., Jaremus, F., & Miller, A. (2022). Do disadvantaged schools have poorer teachers? Rethinking assumptions about the relationship between teaching quality and school-level advantage. *The Australian Educational Researcher*, *49*(4), 635-656. https://doi.org/10.1007/s13384-021-00460-w

Hanushek, E. A., & Woessmann, L. (2017). School resources and student achievement: A review of cross-country economic research. *Cognitive abilities and educational outcomes: A festschrift in honour of Jan-eric Gustafsson*, 149-171. http://dx.doi.org/10.1007/978-3-319-43473-5_8

Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for teaching: Examining personal and institutional predictors of high-quality instruction. *AERA Open, 1*(4), 2332858415617703. https://doi.org/10.1177/2332858415617703

Huang, M. H., & Chang, S. H. (2017). Similarities and differences in East Asian Confucian Culture: A comparative analysis. *OMNES: The Journal of Multicultural Society*, *7*(2), 1-40. http://dx.doi.org/10.15685/omnes.2017.01.7.2.1

Jerrim, J., Oliver, M., & Sims, S. (2022). The relationship between inquiry-based teaching and students' achievement: New evidence from a longitudinal PISA study in England. *Learning and Instruction*, *80*, 101310. https://doi.org/10.1016/j.learninstruc.2020.101310

Koirikivi, P., Benjamin, S., Hietajarvi, L., Kuusisto, A., & Gearon, L. (2021). Resourcing resilience: educational considerations for supporting well-being and preventing violent extremism amongst Finnish youth. *International Journal of Adolescence and Youth, 26*(1), 553-569. https://doi.org/10.1080/02673843.2021.2010578

Luthar, S. S. (2006). Resilience in development: A synthesis of research across five decades.

Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology, 46*(1), 53-83. https://doi.org/10.1016/j.jsp.2007.01.002

Martin, A. J., Burns, E. C., Collie, R. J., Cutmore, M., MacLeod, S., & Donlevy, V. (2022). The role of engagement in immigrant students' academic resilience. *Learning and Instruction*, *82*, 1-9. https://dx.doi.org/10.1016/j.learninstruc.2022.101650

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2019/frameworks/.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/international-results/

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (Eighth ed.). Los Angeles, CA: Muthén & Muthén.

Nagengast, B., & Marsh, H. (2014). Motivation and engagement in science around the globe: Testing measurement invariance with multigroup structural equation models across 57 countries using PISA 2006. In L. Rutkowski, M. von Davier & D. Rutkowski (Eds), *Handbook of international large-scale assessment: Background, technical issues, and data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press, pp. 317-344.

Ni, H., Li, C., Li, B., & Xi, H. (2020). Elementary students' perceptions of classroom resilience-promoting factors in China and the United States. *International Journal of School & Educational Psychology*, *8*(1), 62-73. https://doi.org/10.1080/21683603.2018.1523030

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535-569. https://doi.org/10.1080/10705510701575396

OECD. (2011). *Against the Odds: Disadvantaged Students Who Succeed in School*. Paris: OECD Publishing.

OECD. (2020). Sampling. In *PISA 2018 Technical Report*. Retrieved from https://www.oecd.org/pisa/pisa-for-development/pisafordevelopment2018technicalreport/PISA_D_Chapter_04_Sampling.pdf

Özcan, B., & Bulus, M. (2022). Protective factors associated with academic resilience of adolescents in individualist and collectivist cultures: Evidence from PISA 2018 large scale assessment. *Current Psychology*, *41*(4), 1740-1756. https://doi.org/10.1007/s12144-022-02944-z

Putwain, D. W., & Daly, A. L. (2013). Do clusters of test anxiety and academic buoyancy differentially predict academic performance? *Learning and Individual Differences, 27,* 157-162. https://doi.org/10.1016/j.lindif.2013.07.010

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Sun, J., & Stewart, D. (2007). Age and gender effects on resilience in children and adolescents. *International Journal of mental health promotion*, *9*(4), 16-25. https://doi.org/10.1016/S2542-5196(21)00278-3

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

Tudor, K. E., & Spray, C. M. (2017). Approaches to measuring academic resilience: A systematic review. *International Journal of Research Studies in Education*, *7*(4). https://doi.org/10.5861/ijrse.2017.1880

Vicente, I., Pastor, J. M., & Soler, Á. (2021). Improving educational resilience in the OECD countries: Two convergent paths. *Journal of Policy Modeling*, *43*(6), 1149-1166. https://doi.org/10.1016/j.jpolmod.2021.09.007

Wang, M. C., Deng, Q., Bi, X., Ye, H., & Yang, W. (2017). Performance of the entropy as an index of classification accuracy in latent profile analysis: a monte carlo simulation study. *Acta Psychologica Sinica,* 49(11), 1473-1482. https://doi.org/10.3724/SP.J.1041.2017.01473

Wayman, J. C. (2002). The utility of educational resilience for studying degree attainment in school dropouts. *The Journal of Educational Research*, *95*(3), 167-178. https://doi.org/10.1080/00220670209596587

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: A guide to best practice. *Journal of Black Psychology*, *46*(4), 287-311. https://doi.org/10.1177/0095798420930932

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. versio*n*, *2*(1), 1-189. http://ggplot2.org

Supplementary Materials

Table 1

Highest Educational Level (Raw data vs One-teacher data)

						BTBG04 (Highest Educational Level)				
	AVG	**ONE-T**	**SD**	**RAW**	**SD**	**statistic**	**parameter**	**p_value**	**conf_int_lower**	**conf_int_upper**
AUS	1.49	5.23	0.48	5.22	0.47	0.31	562.60	0.76	-0.06	0.09
BHR	1.98	5.20	0.59	5.21	0.60	-0.16	217.18	0.87	-0.15	0.13
CHL	1.02	5.10	0.33	5.11	0.33	-0.13	278.99	0.90	-0.08	0.07
TWN	1.36	5.60	0.51	5.59	0.51	0.30	416.81	0.76	-0.08	0.11
CYP	1.68	5.69	0.55	5.69	0.52	-0.02	138.77	0.99	-0.16	0.16
FIN	2.15	6.00	0.30	5.91	0.54	2.37	462.95	0.02	0.02	0.17
FRA	1.17	5.80	0.67	5.78	0.66	0.21	247.91	0.84	-0.15	0.18
GEO	1.19	5.79	0.73	5.81	0.67	-0.29	276.58	0.77	-0.18	0.14
HKG	1.27	5.42	0.53	5.40	0.52	0.25	251.19	0.80	-0.11	0.14
HUN	1.64	5.26	0.50	5.27	0.51	-0.14	297.34	0.89	-0.11	0.10
IRN	1.00	5.21	0.58	5.21	0.58	0.02	406.98	0.99	-0.11	0.11
IRL	3.05	5.35	0.59	5.42	0.57	-1.31	226.37	0.19	-0.19	0.04
ITA	1.29	6.07	0.44	6.07	0.41	-0.12	310.99	0.90	-0.10	0.09
JPN	1.11	5.09	0.32	5.08	0.30	0.25	282.56	0.80	-0.06	0.08
KAZ	1.33	5.09	0.50	5.12	0.48	-0.52	314.25	0.61	-0.13	0.08
JOR	1.01	4.76	0.88	4.76	0.87	-0.01	398.98	0.99	-0.17	0.17
KOR	1.22	5.39	0.49	5.39	0.49	-0.15	352.39	0.88	-0.11	0.09
LBN	1.03	5.19	0.84	5.14	0.96	0.62	368.53	0.53	-0.13	0.24
LTU	1.24	5.45	0.60	5.46	0.58	-0.16	382.47	0.88	-0.13	0.11
MYS	1.51	5.03	0.52	5.02	0.48	0.20	334.46	0.84	-0.09	0.11
MAR	1.05	4.00	1.44	3.93	1.47	0.45	365.83	0.65	-0.23	0.37
OMN	1.06	5.03	0.47	5.04	0.47	-0.27	405.98	0.79	-0.10	0.08
NZL	2.42	5.47	0.61	5.50	0.60	-0.39	210.94	0.70	-0.16	0.11
NOR	1.48	5.39	0.54	5.45	0.54	-0.97	243.82	0.33	-0.19	0.07
PRT	1.18	5.18	0.45	5.19	0.45	-0.15	300.50	0.88	-0.11	0.09
QAT	1.39	5.34	0.51	5.28	0.51	0.87	275.20	0.39	-0.06	0.17
ROM	1.09	5.12	0.76	5.13	0.75	-0.07	367.07	0.95	-0.16	0.15
RUS	1.00	5.71	0.52	5.71	0.52	0.00	390.00	1.00	-0.10	0.10
SAU	1.09	5.03	0.16	5.04	0.19	-0.50	319.94	0.62	-0.05	0.03
ZAF	1.01	4.79	0.49	4.78	0.52	0.33	892.52	0.74	-0.05	0.08
SWE	1.40	5.16	0.94	5.22	0.98	-0.52	290.51	0.60	-0.27	0.16
ARE	1.53	5.26	0.56	5.27	0.61	-0.51	1114.31	0.61	-0.08	0.05
TUR	1.03	5.06	0.24	5.06	0.24	0.06	317.57	0.95	-0.05	0.05
EGY	1.02	4.73	0.84	4.72	0.85	0.10	293.99	0.92	-0.18	0.20
USA	1.48	5.57	0.50	5.56	0.50	0.26	515.23	0.79	-0.07	0.09
ENG	1.02	5.29	0.51	5.28	0.51	0.09	151.81	0.93	-0.15	0.17

Note. AVG = Average number of teachers in raw data, SD = standard deviation, RAW = raw data, ONE-T = one teacher data, conf_in = confidence interval.

Table 2

Major (Raw data vs One-teacher data)

						BTDMMME (Major)				
	AVG	ONE-T	SD	RAW	SD	statistic	parameter	p_value	conf_int_lower	conf_int_upper
AUS	1.49	3.78	1.25	3.76	1.26	0.25	561.32	0.80	-0.17	0.22
BHR	1.98	4.33	0.78	4.29	0.80	0.36	210.94	0.72	-0.15	0.22
CHL	1.02	4.33	0.97	4.33	0.96	0.00	274.75	1.00	-0.23	0.23
TWN	1.36	4.05	1.00	3.71	1.17	3.31	451.62	0.00	0.14	0.53
CYP	1.68	4.16	0.63	4.14	0.53	0.20	126.81	0.84	-0.16	0.20
FIN	2.15	3.74	1.02	3.38	1.16	3.41	333.30	0.00	0.15	0.56
FRA	1.17	4.29	0.80	4.27	0.81	0.16	248.20	0.87	-0.18	0.22
GEO	1.19	4.56	1.00	4.60	0.94	-0.25	163.57	0.80	-0.33	0.25
HKG	1.27	3.64	1.14	3.60	1.17	0.25	253.75	0.80	-0.24	0.32
HUN	1.64	3.40	0.76	3.36	0.74	0.45	289.55	0.66	-0.12	0.19
IRN	1.00	3.86	0.98	3.86	0.99	-0.06	406.99	0.95	-0.20	0.19
IRL	3.05	4.05	1.04	3.98	1.07	0.71	235.98	0.48	-0.13	0.28
ITA	1.29	3.65	1.33	3.74	1.31	-0.58	318.26	0.56	-0.37	0.20
JPN	1.11	4.16	0.98	4.17	0.96	-0.09	286.78	0.93	-0.23	0.21
KAZ	1.33	4.36	0.62	4.39	0.60	-0.54	312.30	0.59	-0.17	0.09
JOR	1.01	3.81	0.89	3.80	0.90	0.03	452.00	0.97	-0.16	0.17
KOR	1.22	3.63	0.82	3.68	0.82	-0.51	352.92	0.61	-0.21	0.13
LBN	1.03	3.61	1.13	3.56	1.17	0.49	370.96	0.62	-0.18	0.29
LTU	1.24	4.36	0.73	4.36	0.74	0.10	393.04	0.92	-0.14	0.15
MYS	1.51	3.98	1.12	3.91	1.16	0.62	359.80	0.54	-0.15	0.30
MAR	1.05	3.11	1.45	3.07	1.46	0.32	466.28	0.75	-0.22	0.31
OMN	1.06	4.24	0.75	4.24	0.75	-0.04	431.98	0.97	-0.14	0.14
NZL	2.42	3.46	1.25	3.46	1.23	-0.04	211.97	0.97	-0.28	0.26
NOR	1.48	3.65	0.93	3.70	0.93	-0.50	241.64	0.62	-0.28	0.17
PRT	1.18	4.77	0.60	4.79	0.56	-0.35	296.68	0.72	-0.15	0.11
QAT	1.39	4.35	0.81	4.33	0.80	0.17	274.09	0.87	-0.17	0.20
ROM	1.09	4.21	0.50	4.19	0.51	0.40	312.20	0.69	-0.09	0.14
RUS	1.00	4.66	0.49	4.66	0.49	0.00	390.00	1.00	-0.10	0.10
SAU	1.09	4.08	0.62	4.09	0.63	-0.21	371.05	0.83	-0.14	0.11
ZAF	1.01	4.04	0.93	4.03	0.94	0.18	933.00	0.86	-0.11	0.13
SWE	1.40	3.84	1.30	3.82	1.28	0.19	299.52	0.85	-0.25	0.31
ARE	1.53	4.34	0.83	4.31	0.85	0.70	1070.04	0.49	-0.06	0.13
TUR	1.03	4.26	0.94	4.24	0.93	0.18	313.78	0.86	-0.19	0.23
EGY	1.02	4.10	0.97	4.10	0.97	-0.04	310.79	0.97	-0.22	0.21
USA	1.48	3.57	1.25	3.46	1.25	1.02	515.52	0.31	-0.10	0.31
ENG	1.02	4.32	0.75	4.33	0.75	-0.14	151.88	0.88	-0.26	0.22

Note. AVG = Average number of teachers in raw data, SD = standard deviation, RAW = raw data, ONE-T = one teacher data, conf_in = confidence interval.

Table 3

Cognitive Activation (Raw data vs One-teacher data)

	CAgSCALE (Cognitive Activation)									
	AVG	ONE-T	SD	RAW	SD	statistic	parameter	p_value	conf_int_lower	conf_int_upper
AUS	1.49	22.40	3.30	22.33	3.24	0.25	553.21	0.81	-0.45	0.58
BHR	1.98	25.03	2.81	24.93	2.79	0.30	213.85	0.77	-0.56	0.75
CHL	1.02	24.03	3.16	24.12	3.18	-0.22	270.93	0.82	-0.84	0.67
TWN	1.36	19.31	3.41	19.43	3.44	-0.37	420.03	0.71	-0.75	0.52
CYP	1.68	23.72	2.92	23.63	2.96	0.21	143.01	0.83	-0.80	0.99
FIN	2.15	21.72	3.14	21.27	3.44	1.39	321.06	0.16	-0.18	1.07
FRA	1.17	21.03	2.38	21.04	2.40	-0.01	245.68	0.99	-0.60	0.59
GEO	1.19	23.33	2.60	23.23	2.59	0.34	288.88	0.74	-0.49	0.69
HKG	1.27	19.07	3.34	18.91	3.48	0.38	257.84	0.70	-0.66	0.98
HUN	1.64	22.60	3.03	22.32	3.18	0.83	300.86	0.41	-0.38	0.93
IRN	1.00	22.44	3.29	22.43	3.29	0.04	400.99	0.97	-0.63	0.66
IRL	3.05	21.75	3.45	21.63	3.52	0.34	237.59	0.73	-0.55	0.79
ITA	1.29	23.48	3.34	23.61	3.36	-0.34	321.62	0.73	-0.84	0.59
JPN	1.11	19.29	3.12	19.22	3.20	0.18	285.18	0.85	-0.66	0.80
KAZ	1.33	25.06	2.50	24.83	2.56	0.85	333.77	0.39	-0.30	0.76
JOR	1.01	23.55	2.84	23.56	2.84	-0.05	455.96	0.96	-0.53	0.51
KOR	1.22	20.32	3.81	20.33	3.72	-0.03	344.54	0.98	-0.79	0.77
LBN	1.03	23.07	3.41	23.13	3.40	-0.16	370.66	0.87	-0.75	0.64
LTU	1.24	22.43	2.83	22.35	2.89	0.29	390.80	0.77	-0.48	0.64
MYS	1.51	22.39	3.48	21.93	3.82	1.28	380.23	0.20	-0.25	1.17
MAR	1.05	22.00	3.23	21.88	3.30	0.40	466.67	0.69	-0.47	0.71
OMN	1.06	23.97	2.68	23.94	2.67	0.13	439.89	0.90	-0.47	0.53
NZL	2.42	22.68	3.55	22.64	3.35	0.09	205.68	0.93	-0.72	0.79
NOR	1.48	20.99	3.00	21.15	3.16	-0.41	224.41	0.68	-0.93	0.61
PRT	1.18	22.18	3.08	21.89	3.41	0.80	308.75	0.42	-0.43	1.02
QAT	1.39	24.47	3.09	24.44	3.11	0.10	272.81	0.92	-0.67	0.74
ROM	1.09	23.37	3.19	23.40	3.26	-0.07	371.17	0.94	-0.68	0.63
RUS	1.00	22.66	2.97	22.66	2.97	0.00	390.00	1.00	-0.59	0.59
SAU	1.09	24.04	3.04	24.06	3.10	-0.07	346.80	0.94	-0.67	0.62
ZAF	1.01	21.97	3.66	21.95	3.65	0.07	938.83	0.94	-0.45	0.48
SWE	1.40	21.55	3.15	21.83	3.33	-0.78	302.12	0.44	-0.99	0.43
ARE	1.53	25.04	2.95	24.97	3.01	0.41	1031.62	0.68	-0.27	0.42
TUR	1.03	23.64	3.15	23.65	3.18	-0.03	315.93	0.98	-0.71	0.69
EGY	1.02	23.09	2.96	23.09	2.95	0.01	322.85	1.00	-0.64	0.65
USA	1.48	24.02	3.03	23.78	3.16	0.91	515.89	0.36	-0.27	0.75
ENG	1.02	21.70	3.06	21.73	3.05	-0.04	141.86	0.97	-1.03	0.99

Note. AVG = Average number of teachers in raw data, SD = standard deviation, RAW = raw data, ONE-T = one teacher data, conf_in = confidence interval.

Table 4

Teachers' Ratings on School Emphasis on Academic Success (Raw data vs One-teacher data)

	BTBGEAS (school emphasis on academic success: teachers' perspective)									
	AVG	**ONE-T**	**SD**	**RAW**	**SD**	**statistic**	**parameter**	**p_value**	**conf_int_lower**	**conf_int_upper**
AUS	1.49	9.91	2.11	9.94	2.07	-0.22	559.92	0.82	-0.37	0.29
BHR	1.98	11.03	1.79	10.87	1.87	0.73	221.64	0.47	-0.27	0.58
CHL	1.02	9.68	2.38	9.69	2.56	-0.03	278.19	0.98	-0.59	0.57
TWN	1.36	9.95	2.03	9.73	2.00	1.17	418.82	0.24	-0.15	0.59
CYP	1.68	10.82	2.61	10.57	2.48	0.64	140.01	0.52	-0.52	1.02
FIN	2.15	9.59	1.66	9.45	1.55	0.87	278.75	0.39	-0.18	0.45
FRA	1.17	9.29	1.56	9.30	1.58	-0.07	250.00	0.95	-0.40	0.37
GEO	1.19	9.92	1.63	9.89	1.56	0.19	288.72	0.85	-0.33	0.40
HKG	1.27	9.48	1.81	9.45	1.76	0.14	249.93	0.89	-0.40	0.46
HUN	1.64	9.22	1.70	9.27	1.61	-0.29	279.77	0.77	-0.41	0.30
IRN	1.00	9.48	2.25	9.47	2.25	0.05	408.99	0.96	-0.43	0.45
IRL	3.05	10.70	1.87	10.64	1.95	0.30	234.46	0.77	-0.31	0.43
ITA	1.29	9.38	1.52	9.45	1.56	-0.38	323.93	0.70	-0.39	0.26
JPN	1.11	9.65	1.93	9.59	1.91	0.27	286.71	0.79	-0.38	0.50
KAZ	1.33	11.45	1.50	11.23	1.59	1.34	345.64	0.18	-0.10	0.54
JOR	1.01	9.79	1.92	9.80	1.93	-0.10	459.98	0.92	-0.37	0.33
KOR	1.22	11.33	2.44	11.33	2.47	-0.01	353.30	0.99	-0.51	0.50
LBN	1.03	9.82	2.10	9.86	2.11	-0.22	384.81	0.83	-0.47	0.37
LTU	1.24	10.35	1.27	10.31	1.32	0.32	396.30	0.75	-0.21	0.29
MYS	1.51	10.76	1.95	10.56	1.79	1.06	336.59	0.29	-0.17	0.57
MAR	1.05	8.42	1.77	8.40	1.78	0.15	482.21	0.88	-0.29	0.34
OMN	1.06	10.59	1.74	10.60	1.76	-0.07	438.94	0.95	-0.34	0.32
NZL	2.42	10.20	2.35	10.19	2.21	0.03	199.50	0.98	-0.50	0.51
NOR	1.48	9.99	1.56	9.95	1.49	0.21	233.07	0.83	-0.33	0.41
PRT	1.18	9.17	1.69	9.15	1.64	0.08	300.94	0.94	-0.36	0.39
QAT	1.39	11.19	2.08	11.19	2.08	0.00	276.34	1.00	-0.47	0.47
ROM	1.09	10.18	1.86	10.16	1.88	0.14	372.55	0.89	-0.35	0.41
RUS	1.00	9.46	1.37	9.46	1.37	0.00	390.00	1.00	-0.27	0.27
SAU	1.09	11.25	2.20	11.23	2.19	0.09	384.89	0.93	-0.42	0.46
ZAF	1.01	9.25	2.09	9.26	2.08	-0.08	954.85	0.94	-0.27	0.25
SWE	1.40	9.67	1.56	9.65	1.64	0.14	309.74	0.89	-0.32	0.37
ARE	1.53	11.45	2.51	11.52	2.57	-0.47	1076.65	0.64	-0.36	0.22
TUR	1.03	9.49	2.06	9.48	2.07	0.05	317.85	0.96	-0.44	0.47
EGY	1.02	10.16	1.84	10.19	1.83	-0.13	320.84	0.90	-0.43	0.38
USA	1.48	9.63	2.16	9.64	2.12	-0.07	509.49	0.95	-0.36	0.34
ENG	1.02	10.63	2.07	10.66	2.09	-0.09	151.96	0.92	-0.69	0.63

Note. AVG = Average number of teachers in raw data, SD = standard deviation, RAW = raw data, ONE-T = one teacher data, conf_in = confidence interval.

Table 5

Teachers' Ratings on Safe and Orderly Environment (Raw data vs One-teacher data)

	BTBGSOS (Safe and Orderly Environment: teachers' perspective)									
	AVG	**ONE-T**	**SD**	**RAW**	**SD**	**statistic**	**parameter**	**p_value**	**conf_int_lower**	**conf_int_upper**
AUS	1.49	10.59	2.31	10.53	2.28	0.31	559.94	0.76	-0.30	0.42
BHR	1.98	10.90	1.75	10.83	1.84	0.35	226.19	0.72	-0.34	0.49
CHL	1.02	9.25	2.19	9.26	2.22	-0.03	278.97	0.97	-0.53	0.51
TWN	1.36	10.19	1.76	10.10	1.73	0.54	416.40	0.59	-0.24	0.41
CYP	1.68	10.39	2.49	10.16	2.47	0.60	143.40	0.55	-0.53	0.98
FIN	2.15	9.38	1.76	9.26	1.63	0.76	277.11	0.45	-0.20	0.46
FRA	1.17	9.40	1.92	9.49	1.93	-0.38	249.18	0.70	-0.57	0.38
GEO	1.19	11.06	1.53	11.01	1.48	0.27	287.92	0.78	-0.30	0.39
HKG	1.27	10.47	1.93	10.59	1.92	-0.49	250.99	0.62	-0.58	0.35
HUN	1.64	10.05	1.67	10.09	1.67	-0.19	291.47	0.85	-0.39	0.32
IRN	1.00	10.65	2.01	10.65	2.01	0.04	408.99	0.97	-0.38	0.40
IRL	3.05	11.18	2.02	10.86	2.12	1.61	236.18	0.11	-0.07	0.73
ITA	1.29	9.41	1.50	9.36	1.44	0.31	313.93	0.76	-0.27	0.36
JPN	1.11	8.87	1.42	8.86	1.45	0.06	285.19	0.95	-0.32	0.34
KAZ	1.33	11.61	1.76	11.56	1.72	0.25	328.64	0.80	-0.32	0.41
JOR	1.01	10.28	2.00	10.29	2.00	-0.06	443.95	0.95	-0.38	0.36
KOR	1.22	9.44	1.82	9.50	1.81	-0.35	349.04	0.72	-0.44	0.31
LBN	1.03	10.21	1.95	10.26	1.98	-0.25	372.97	0.80	-0.45	0.35
LTU	1.24	10.61	1.67	10.48	1.66	0.76	386.77	0.44	-0.20	0.45
MYS	1.51	9.95	1.84	9.76	1.77	1.08	348.60	0.28	-0.16	0.55
MAR	1.05	9.74	2.15	9.73	2.15	0.06	475.91	0.95	-0.37	0.40
OMN	1.06	11.02	1.63	11.02	1.65	0.03	436.00	0.98	-0.30	0.31
NZL	2.42	10.31	2.18	10.42	2.21	-0.47	216.20	0.64	-0.59	0.36
NOR	1.48	10.47	1.83	10.35	1.82	0.53	239.10	0.60	-0.32	0.56
PRT	1.18	10.22	1.97	10.09	1.97	0.61	303.47	0.54	-0.30	0.57
QAT	1.39	11.04	2.02	11.10	2.06	-0.26	279.36	0.80	-0.52	0.40
ROM	1.09	11.40	2.04	11.38	2.05	0.08	368.88	0.93	-0.40	0.43
RUS	1.00	10.49	1.88	10.49	1.88	0.00	390.00	1.00	-0.37	0.37
SAU	1.09	11.56	1.91	11.54	1.93	0.09	371.68	0.93	-0.37	0.41
ZAF	1.01	8.44	2.01	8.45	2.02	-0.10	960.96	0.92	-0.27	0.24
SWE	1.40	9.44	1.32	9.47	1.34	-0.19	303.78	0.85	-0.32	0.26
ARE	1.53	11.64	2.18	11.67	2.15	-0.29	1043.72	0.77	-0.28	0.21
TUR	1.03	10.03	2.23	9.98	2.23	0.21	317.79	0.83	-0.44	0.54
EGY	1.02	10.91	1.82	10.94	1.83	-0.15	322.90	0.88	-0.43	0.37
USA	1.48	9.81	2.25	9.79	2.28	0.10	515.35	0.92	-0.35	0.39
ENG	1.02	10.59	2.20	10.63	2.20	-0.12	151.91	0.91	-0.74	0.66

Note. AVG = Average number of teachers in raw data, SD = standard deviation, RAW = raw data, ONE-T = one teacher data, conf_in = confidence interval.

Table 6

Number of Disadvantaged Students and Thresholds for Academic Resilience

	Students (*N*)	Students with HER information (*n*)	HER missing rate (%)	Disadvantaged students (*k*)	Thresholds for academic performance
AUS	9060	8902	1.74%	2967	554.28
BHR	5725	5684	0.72%	1895	527.66
CHL	4115	4070	1.09%	1357	471.92
TWN	4915	4910	0.10%	1637	659.63
CYP	3521	3500	0.60%	1167	537.99
FIN	4874	4821	1.09%	1607	540.83
FRA	3874	3724	3.87%	1241	513.02
GEO	3315	3264	1.54%	1088	498.65
HKG	3265	3248	0.52%	1083	620.36
HUN	4569	4531	0.83%	1510	555.47
IRN	5980	5970	0.17%	1990	483.45
IRL	4117	4060	1.38%	1353	557.07
ITA	3619	3605	0.39%	1202	529.04
JPN	4446	4438	0.18%	1479	631.15
JOR	7176	7078	1.37%	2359	459.87
KAZ	4453	4435	0.40%	1478	520.50
KOR	3861	3856	0.13%	1285	651.13
LBN	4730	4667	1.33%	1556	458.89
LTU	3826	3652	4.55%	1217	555.83
MYS	7065	7033	0.45%	2344	495.17
MAR	8458	8383	0.89%	2794	414.10
OMN	6751	6646	1.56%	2215	456.15
NZL	6050	5926	2.05%	1975	517.97
NOR	4575	4270	6.67%	1423	538.54
PRT	3377	3354	0.68%	1118	530.85
QAT	3884	3835	1.26%	1278	483.00
ROM	4494	4440	1.20%	1480	525.09
RUS	3901	3893	0.21%	1298	577.87
SAU	5680	5580	1.76%	1860	425.14
ZAF	20829	20622	0.99%	6874	414.08
SWE	3996	3891	2.63%	1297	536.74
ARE	22334	21778	2.49%	7259	520.74
TUR	4077	4044	0.81%	1348	542.60
EGY	7210	7055	2.15%	2352	455.04
USA	8698	8321	4.33%	2774	561.88
ENG	3365	3183	5.41%	1061	552.35

Note. HER = Home Educational Resources Index, students with the bottom 1/3 of HER were identified as disadvantaged, and performance thresholds were applied to their five plausible values in mathematics to define resilience. Thresholds for performance were calculated based on five plausible values via *intsvy* package in R software.