

The Moderating Role of Teacher Autonomy  
for Classroom Practices and Student Achievement:

Cross-country Analysis

by

Ketevan Chachkhiani

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

Approved June 2023 by the  
Graduate Supervisory Committee:

Margarita Pivovarova, Chair  
David Berliner  
Iveta Silova

ARIZONA STATE UNIVERSITY

August 2023

## ABSTRACT

Autonomy is often discussed as a necessary condition for professionalism in any field, including the teaching profession. In the educational context, autonomy is seen to be critical for teachers to be able to synthesize their knowledge and skills and translate them into effective classroom practices in the best interest of their students. In response to national and global trends primarily associated with neoliberal reforms and their constraining consequences for teacher autonomy, researchers have been trying to unpack this concept and demonstrate evidence on how it affects teachers, students, and the educational system in general. However, this empirical evidence is both quite scarce and controversial. In this dissertation, I define teacher autonomy as a decision-making space created through the freedom from input and output control and identify four types of autonomy: limited, moderate output-driven, moderate input-driven, and extended. Using the data from 43 countries from the Progress in International Reading Literacy Study (PIRLS) 2016 database, I explore the moderating role of each type of autonomy for classroom practices, student achievement, and educational equity. I find that although students in countries with moderate input-driven and extended autonomy have the highest achievement compared to limited and moderate output-driven ones, their teachers, overall, use key classroom practices less frequently, as well as the effectiveness of those practices is relatively low. Findings are more consistent in relation to educational equity. These two groups of countries experience the lowest achievement gap in reading, as well as have teachers who use effective strategies with more frequency with disadvantaged students. In addition, classroom practices in countries with extended teacher autonomy

and, to a certain extent, with moderate input-driven one show more potential to contribute to narrowing the achievement gap.

## DEDICATION

I dedicate my dissertation to my parents, who have always wholeheartedly believed in the power of education. My late dad, Alexander, was anxiously awaiting the decision of the selection committee from Arizona State University but unexpectedly passed away two weeks before the news about my acceptance to the Ph.D. program came. I know he would have been very happy and proud now. My mom, Eteri, has always believed in me and strongly supported and cheered me in every small and big aspiration since I was a child. I know she is very proud and happy now. I thank both for fostering my development with love and care into who I am today.

## ACKNOWLEDGMENTS

A couple of months ago, when sharing my doctoral experience with my former advisor from Columbia University, she told me how lucky and blessed I was to be surrounded and supported by such fantastic scholars at Arizona State University. I truly am. I would not have been able to complete this journey without them by my side.

I cannot find enough words, neither in English nor in Georgian, my native language, to express my heartfelt gratitude for everything Dr. Margarita Pivovarova, my dissertation committee chair, has done for me since our first email exchange five years ago. I feel honored and lucky to have had her as my mentor and a primary advisor throughout my doctoral journey. I am thankful to her for making me part of the fascinating scholarly research and evaluation projects she has led, and for always supporting my research initiatives. This dissertation would not have been possible without her encouragement, dedicated support, and innumerable meetings, phone conversations, emails, and texts from the design stage through the end. I will be forever grateful for her being not only a brilliant professor, scholar, and mentor but also a true and devoted friend who has always been there for me, including during some of the most difficult times of my life.

I also extend my sincere gratitude and appreciation to Dr. Iveta Silova, a member of my dissertation committee and my advisor, who is the reason I decided to pursue my doctoral studies at Mary Lou Fulton Teachers College. Her feedback in framing my future dissertation has been very helpful. I also appreciate her enriching my doctoral experience by engaging me in intriguing research projects and publications I am proud of. I am also grateful for her giving me an opportunity to engage in the efforts to extend

ASU's engagement in improving the teacher education programs in Georgia, my home country. Dr. Silova's excellent guidance, academic and personal support, encouragement, and trust have had a formative role in my doctoral journey.

I am truly privileged to have Dr. David Berliner, Regent's Professor Emeritus and an internationally renowned scholar, on my dissertation committee. I feel honored to have an opportunity to sit in the classes or attend meetings he was invited to speak and share his research findings and perspectives about pressing educational issues in the US and beyond. Also, the feedback I received from Dr. Berliner as part of my comprehensive exams and proposal defense significantly influenced how I have shaped the main research questions of my doctoral study. In addition, I would like to express my appreciation for his always supportive and motivating feedback and communication, which immensely helped me build confidence in my academic work.

I would also like to acknowledge that my doctoral studies at ASU would not be possible without financial support from Mary Lou Fulton Teachers College and Graduate College. Their generous funding for doctoral students, including international ones, has enabled me to work with fantastic scholars and pursue my academic and research goals.

Finally, I am thankful to my wonderful family. First, to my husband, Shalva, who has always supported my personal, educational, and professional goals, including this biggest academic adventure of my life. Finally, I owe a lot to my sons, Dachi and Rati, who followed me from Georgia on my journey to Arizona. Seeing them grow into intelligent, hard-working, passionate, and kindhearted young men has given me the strength to try even harder to complete this pivotal chapter in my life successfully.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	viii
LIST OF FIGURES.....	x
CHAPTER	
1 INTRODUCTION .....	1
Problem and Significance .....	2
Purpose and Research Questions .....	4
Organization of the Dissertation .....	5
2 LITERATURE REVIEW .....	6
Introduction .....	6
Historical and Contextual Overview .....	7
Existing Contextualizations of Teacher Autonomy .....	14
Measuring Teacher Autonomy.....	29
Autonomy as a Source of Professionalism.....	32
Teacher Autonomy Enhancing Student Learning and Equity .....	34
Teacher Autonomy Creating Risks for Student Learning and Equity.....	39
Autonomy – Suitable for all Teachers? .....	46
The Impact of Teacher Autonomy on Student and Achievement and Achievement Gap – Existing Empirical Evidence .....	51
3 METHODOLOGY .....	55
Data .....	55
Participants and Sampling.....	57

CHAPTER	Page
Measures .....	62
Data Analysis .....	81
4 RESULTS .....	89
Descriptive Statistics .....	89
Main Findings.....	92
5 DISCUSSION AND CONCLUSION .....	176
Summary of Findings .....	176
Discussion and Implications .....	193
Study Limitations .....	211
Conclusion .....	212
REFERENCES .....	213
APPENDIX	
A SUMMARY OF MEASURES .....	235



## LIST OF TABLES

Table	Page
1. PIRLS 2016 Countries Selected for the Study .....	58
2. Items for Operationalizing Teacher Autonomy as Freedom from Control .....	63
3. The Description for Input and Output Control Score Adjustment .....	65
4. Steps of Entering Variables in Linear Regression .....	86
5. Descriptive Statistics.....	91
6. PIRLS 2016 Countries: Scores and Types of Input and Output Control .....	97
7. Countries by Autonomy, Economic and Democratic Development .....	99
8. Student Achievement by Teacher Autonomy .....	104
9. Achievement Gap by Teacher Autonomy .....	109
10. Classroom Practices by Teacher Autonomy .....	113
11. Individualized Instruction by Teacher Autonomy .....	113
12. Classroom Practices by the Share of Disadvantaged Students .....	120
13. Individualized Instruction by the Share of Disadvantaged Students .....	121
14. Classroom Practices by Teachers' Perceptions .....	127
15. Individualized Instruction by Teacher's Perceptions .....	129
16. Teacher Qualification Characteristics by Teacher Autonomy .....	132
17. Autonomy Moderating the Effect of Teaching Experience on Reading Comprehension Strategies .....	136
18. Autonomy Moderating the Effect of Teaching Experience on Reading Content Differentiation .....	137

Table	Page
19. Autonomy Moderating the Effect of Teaching Experience on Autonomy-supportive Strategies .....	139
20. Autonomy Moderating the Effect of School-based Professional Development on Reading Comprehension Strategies.....	141
21. Autonomy Moderating the Effect of School-based Professional Development on Reading Content Differentiation .....	143
22. Autonomy Moderating the Effect of School-based Professional Development on Autonomy-Supportive Strategies .....	145
23. Intercorrelations Results for Classroom Practices .....	148
24. Autonomy as a Moderator: Reading Comprehension Strategies and Student Achievement .....	151
25. Autonomy as a Moderator: Reading Content Differentiation and Student Achievement .....	153
26. Autonomy as a Moderator: Autonomy-supportive Strategies and Student Achievement .....	155
27. Autonomy as a Moderator: Individualized Instruction and Student Achievement .....	157
28. Autonomy Moderating the Effect of Reading Comprehension Strategies .....	164
29. Autonomy Moderating the Effect of Reading Content Differentiation .....	167
30. Autonomy Moderating the Effect of Autonomy-supportive Strategies .....	170
31. Autonomy Moderating the Effect of Individualized Instruction .....	173

## LIST OF FIGURES

Figure	Page
1. Teacher Autonomy: Summarizing the Measure .....	68
2. Distribution of Countries by Teacher Autonomy .....	95
3. Achievement Gap by Achievement Quartiles .....	105
4. Achievement Gap by Student Socioeconomic Background .....	107
5. Achievement Gap by the Share of Disadvantaged Students.....	107
6. Reading Comprehension Strategies by the Share of Disadvantaged Students ..	115
7. Reading Content Differentiation by the Share of Disadvantaged Students .....	117
8. Autonomy-supportive Strategies by the Share of Disadvantaged Students.....	119
9. Reading Comprehension Strategies by Teachers' Perceptions .....	123
10. Reading Content Differentiation by Teachers' Perceptions .....	125
11. Autonomy-supportive Strategies by Teachers' Perceptions .....	126
12. An Analysis Model for Research Question 4 .....	131
13. An Analysis Model for Research Question 5 .....	148
14. An Analysis Model for Research Question 6 .....	160

## CHAPTER 1

### INTRODUCTION

Teacher autonomy has become one of the captivating terms in contemporary educational research and practice. Scholarly and policy-level debates about teacher autonomy have expanded since the mid-1980s, when the introduction of neoliberal principles affected all domains of life, including national educational systems and the teaching profession (Codd, 2005; Robertson, 2008). Along with placing education within the “orbit of economic policy” (Codd, 2005, p. 193), schools were tasked to focus on preparing workforce for the economy and equip them with the necessary education and skills “to compete efficiently and effectively” (Manteaw, 2008, p. 121). Scholars argue that such a purely economic emphasis on the purpose of public education turned teachers into skilled technicians, or managed professionals, with narrowly described technical competencies and the obligation of producing pre-determined outcomes (Day, 2002). As autonomy was often seen as a source of professionalism in any field, including education, researchers started to express concerns about the implications of restricted freedom for teacher lives, classroom behavior, and, most importantly, student achievement and equity (Berliner & Glass, 2014; Webb, 2002).

Although neoliberal principles have been shaping individual countries’ educational policies worldwide, they are not always coherent and similar across nations. These policies “are never simply implemented but are translated, mediated and enacted within specific socio-material contexts in which interacting discourses, practices and actors compete” (Gobby et al., 2018, p. 162). The dominant rationalities of neoliberal policies are also contested by the traditional understanding of professionalism and its

relationship to democracy and social and equity goals (Gobby et al., 2018). As a result, the effect of neoliberal reforms on teacher work and autonomy varies by region and country, given the differences in the nature of liberal reforms, as well as enabling and constraining contextual factors affecting teacher autonomy (Ashley, 2012; Erss et al., 2016; MacDonald, 2002; Mausethagen & Molstad 2015; Paradis et al., 2019; Silova et al., 2021).

In this dissertation, I aim to unpack how different countries across the world have shaped teacher autonomy, defined as a decision-making space (Luthans, 1992, seen in Friedman, 1999) and “freedom from external interference, pressure and control” (Charterts, 1976, as cited in Strong & Yoshida, 2014, p. 127) in the domain of curriculum implementation and assessment. More importantly, I explore how different types of autonomy, considering the extent of freedom from government-imposed input and output control, manifest in teacher classroom behaviors and eventually in student achievement and educational equity. In this analysis, I use the data from the 2016 Progress in International Reading Literacy Study (PIRLS) from 43 countries worldwide.

### **Problem and Significance**

In response to national and global trends primarily associated with neoliberal reforms and their consequences, autonomy emerges as a critical variable when examining educational reform initiatives. Researchers have been trying to unpack autonomy and demonstrate evidence on how it affects teachers, students, and the educational system in general. Findings and positions researchers take toward autonomy are quite contradictory. On the one hand, researchers argue that granting autonomy and empowering teachers is an appropriate place to begin solving the problems of today's schools (Pearson &

Moomaw, 2005). Autonomy is believed to be the main condition for teacher well-being and job satisfaction (Avanzi et al., 2013; Berry et al., 2010) and a source of professionalism (Biesta, 2015a, Freidson, 2001) and meaningful and active professional learning (Hargreaves et al., 2013). Other pieces of literature also claim that teacher autonomy positively affects student learning (Hyslop-Margison & Sears, 2010; Pearson & Hall, 1993; Prichard & Moore, 2016), as well as serves as a source of effectiveness for the educational system (Cribb & Gewirtz, 2017; Guarino et al., 2006).

On the other hand, we cannot overlook dangers associated with autonomy, such as potential risks to equity (Boote, 2006; LaCoe, 2008; Murphy et al., 1986), maintenance of the status quo (Erss, 2018; Johnson, 2016) and potential for opportunistic behavior (Woessmann, 2005), personal costs for teachers (Bakken, 2019; Benson, 2010; Ronnberg, 2007), and barriers it may create to teacher collaboration (Anderson, 1987; Bakken, 2019). Most importantly, researchers express concerns about its suitability for all types of teachers and their ability to exercise it as intended for the best interests of students (Boote, 2006; Fox, 2012; Grant et al., 2020).

Understanding the implications of empirical studies and conceptual pieces for policy-making purposes is complicated by the inconsistent conceptualizations of teacher autonomy. As Gwaltney (2012) points out, one needs to make many efforts to understand whether researchers are capturing the entire concept of autonomy, just one aspect of it, or something different. In addition to differences in conceptualizations (for example, teacher autonomy as a decision-making space or a self-governance capacity), existing measures of teacher autonomy do not sufficiently represent the relational and multifaceted nature of teacher autonomy. For example, no empirical studies explicitly account for input and

output control mechanisms when conceptualizing teacher autonomy as a decision-making space.

Despite an ample interest in teacher autonomy, cross-country research on the implications of teacher autonomy for teacher-classroom performance and student outcomes is limited. This gap further widens as we try to unpack these implications for teachers with different qualification characteristics. As such, the study contributes to the body of literature on teacher autonomy in two ways. First, in this study, I use a more comprehensive construct of teacher autonomy, as a freedom from control, by accounting for the presence or absence of both input and output control mechanisms. Second, the findings and implications of this study add to the existing scarce literature on the implications of teacher autonomy for teachers' classroom practices, student achievement, and equity in outcomes.

### **Purpose and Research Questions**

In this dissertation, I explore the moderating role of teacher autonomy for teachers' classroom practices and student achievement in reading. Specifically, I address the following research questions:

(1) How are PIRLS 2016 participant countries distributed by types of teacher autonomy in curriculum implementation and assessment?

(2) How does a) student achievement in reading and b) the achievement gap in reading compare across the national systems with different types of teacher autonomy?

(3) How does the use of teacher classroom practices, such as a) reading comprehension strategies, b) reading content differentiation, c) autonomy-supportive

strategies, and d) individualized instruction, compare across the national systems with different types of teacher autonomy?

(4) How does teacher autonomy moderate the relationship between teacher qualification characteristics (teaching experience and school-based professional development) and the use of classroom practices (reading comprehension strategies, reading content differentiation, autonomy-supportive strategies)?

(5) How does teacher autonomy moderate the relationship between classroom practices (reading comprehension strategies, reading content differentiation, autonomy-supportive strategies, and individualized instruction) and student achievement in reading?

(6) How does teacher autonomy affect the influence of teacher practices (reading comprehension strategies, reading content differentiation, autonomy-supportive strategies, and individualized instruction) on the relationship between student socioeconomic background and student achievement in reading?

### **Organization of the Dissertation**

I begin in Chapter 2 with a review of the literature in which I present the historical overview of the development of teacher autonomy, major conceptualizations of its multifaceted nature, and existing measures. Most importantly, I discuss two bodies of scholarly literature supporting and criticizing broad autonomy for teachers. In Chapter 3, I describe the methodology of the study. Specifically, I discuss the source of data and participants, measures, and statistical analytical tools used for the study. In Chapter 4, I present the findings of the study for all six research questions. In concluding Chapter 5, I summarize the findings for each type of teacher, followed by the discussion and implications of the results.



## CHAPTER 2

### LITERATURE REVIEW

#### **Introduction**

In this chapter, I review scholarly literature related to the key concept of this study, teacher autonomy. First, I present the historical overview of the development teacher autonomy has undergone since the 1960s, with particular attention to the effects of neoliberal reforms on teacher autonomy across different parts of the world. Next, I summarize the major conceptualizations of teacher autonomy within the existing scholarly literature. Specifically, I draw differences between two approaches to this concept, teacher autonomy as a decision-making space and self-governance capacity. Based on Cribb and Gewirtz's framework, I also overview the multifaceted nature of teacher autonomy and its three major dimensions: loci and modes of autonomy, domains of autonomy, and loci and modes of control. Then I present existing measures of teacher autonomy, developed and used by the schools to study the concept of autonomy over the years. To unpack the contradictions about the implications of teacher autonomy for student learning and achievement, I highlight two major bodies of literature. First, I present the findings and conceptual arguments claiming the potential of teacher autonomy to positively impact student learning and achievement through various mechanisms. Second, I discuss the threats scholars identify concerning teacher autonomy. Considering potential risks, I also discuss existing scholarly contributions on the suitability of teacher autonomy for all teachers. I conclude the chapter by highlighting empirical findings on the relationship between teacher autonomy and student achievement.

## **Historical and Contextual Overview**

Approaches to teacher professionalism and autonomy have undergone drastic changes since the 1960s. Through the mid-1970s, the period can be defined as the area of the welfare state, with the so-called "activist government" (Basu, 2004, p. 622) and the confidence in the state to solve all the social problems (Hargreaves & Shirley, 2009). Along with expanding public services, the state generously funded the public education system (Basu, 2004). It based this decision on the belief that education was critical for maintaining social democracies through its power to raise informed and educated citizens to participate and contribute to society (Davies & Bansel, 2007). The state hired teachers to play a critical role in fulfilling this goal. As a result, teaching status significantly improved in this period, reflecting increased salaries and teacher education programs. Mostly, teachers enjoyed broad autonomy over curriculum development and classroom implementation and were not constrained by external examinations or other controlling factors (Codd, 2005). As Hargreaves (2000) points out, the level of autonomy was so high that teaching started to resemble a fully isolated activity behind closed doors.

The area of unlimited teacher autonomy ended in the mid-1980s due to the adoption of neoliberal policies in all domains of life (Codd, 2005). The transition from social welfare to a market-driven economy had strong implications for education. Robertson (2008) identifies three common dimensions of change in the attempts to restructure the system of education worldwide according to neoliberalism principles. The first principle is related to the mandate, i.e., the purpose of the educational system. The economy became the highest priority and the key factor determining the state's involvement in education (Robertson, 2008). Education was seen as a condition for

economic growth, tasked to train the future workforce for the economy and give them the necessary education and skills "to compete efficiently and effectively" (Manteaw, 2008, p. 121; Robertson, 2008). The second principle introduced a new vision about the forms of capacity. i.e., means that can be used for realizing the mandate. Educational organizations were tasked with using their significantly decreased efficiency and generating new funds themselves (Robertson, 2008). Finally, neoliberal policies introduced new mechanisms of governing and coordinating the system. The new model of governance, known as New Public Management, included such elements as putting education providers into competition with each other, decentralizing some services while keeping others centralized, and attaching funding to performance and outcomes (Day, 2002; Robertson, 2008).

Curriculum standardization was one of the significant neoliberal reforms in education worldwide, guided by the new principle for governing and coordinating the educational system. Education Reform Act, approved by Margaret Thatcher in 1988, among other reform areas, included the development of the national curriculum. Not surprisingly, the standards under this curriculum were developed with the participation of the business sector representatives to ensure the alignment of the labor market needs and the education system (Verger et al., 2016). Similarly, by 2001 almost all US states had developed curriculum standards and standardized tests for their schools (Hursh, 2005). As a result of this reform, teachers were no more responsible for developing curricula. Instead, district or state-level officials, outside experts, and curriculum designers took this task (Wills & Sandholtz, 2009). As Smyth (2001) argues, curriculum standardization, along with tests, appraisal systems, incentives, and punishments for demonstrated

performance, was a recentralization of the control over education while creating the false impression of decentralization and increased power held by those at the local sites.

Increasing emphasis on marketization, managerialism, and competitive performativity significantly changed national education systems and the teaching profession (Gobby et al., 2018). Purely economic emphasis on the purpose of public education turned teachers into skilled technicians, or managed professionals, with narrowly described technical competencies and the obligation of producing pre-determined outcomes (Codd, 2005). Their work increasingly depended on the external definitions of quality (Day, 2002). Critical reflection and professional judgment exercised by a teacher as an autonomous professional were no longer considered valuable actions of teachers (Codd, 2005). The reforms significantly threatened teacher agency and implicitly encouraged them to follow the plans without critical examination (Day, 2002).

Even though neoliberal principles have been shaping national educational policies worldwide, they have had different timing and have not always been coherent and similar. The United States and the United Kingdom were the first to fully replace the welfare state with a market-driven economy. Countries in the southern hemisphere followed them soon and diligently installed all the principles of neoliberalism in all domains of their lives (Davies & Bansel, 2007). Some countries, such as the former Soviet republics, were introduced to neoliberalism later as part of the development assistance (Takala & Piattoeva, 2012) and, similar to others, chose diverse paths for its implementation (Silova, 2018). In addition to different timing, the dominant rationalities of neoliberal policies have also been contested by the traditional understanding of professionalism and its relationship to local context and nations' social and equity goals

(Gobby et al., 2018). As a result, the effect of neoliberal reforms on teacher work and autonomy varies by regions and countries.

Existing research suggests that teacher autonomy and agency issues are particularly problematic in the USA, Great Britain, New Zealand, Australia, and Canada, i.e., countries that most willingly introduced neoliberalism in their educational systems. Standardization and performativity turned teachers in the US into the transmitters of scripted lessons, which school districts often purchase, particularly in core subjects (Wills & Sandholtz, 2009). Under such a scripted curriculum, teacher knowledge, experience, and judgment lose their importance. Crocco and Costigan (2007) found teachers in New York believed that the curriculum standardization and the testing system were undermining all the control they could have over the teaching process and their professional development. Like US colleagues, teachers in Canada have experienced a history of accountability and distrust (Paradis et al., 2018) and have expressed concerns about being treated as technicians rather than professionals (Paradis et al., 2019).

The presence of performativity culture and output control has significantly constrained teacher autonomy and agency, even when the government softens input control. For example, Scotland's "Curriculum for Excellence" policy in 2006 positioned teachers as agents of change and curriculum makers instead of curriculum deliverers (Priestley et al., 2015a). Cyprus took the same approach towards teacher professional development. It replaced the discourse of state-initiated and centralized professional development with one of self-education and self-development (Priestley et al., 2015b). These initiatives are a return to freedom, bottom-up approaches, and shifting away from prescriptive and teacher-proof curricula. However, they were strongly contested by

accountability systems and extensive use of metrics for educational effectiveness (Priestley et al., 2015b).

Implementation of neoliberal reforms and their implications for teacher autonomy has been quite different in Nordic countries, such as Norway, Sweden, Finland, and Denmark. Compared to Anglo-American counterparts, the degree of external control in Nordic countries has been much lower (Mausethagen & Molstad, 2015; Stephens et al., 2004). Also, emphasis on student outcomes is made not only from the transparency and efficiency perspective but also from equality and quality (Mausethagen, 2013). Nevertheless, there are significant differences among the countries as well. All countries have centralized, national-level curriculum documents; however, the level of mandating is different. There are differences in accountability measures and administration of centralized tests as well (Carlgren & Klette, 2008). Erss and Kalmus (2018) also emphasize the changes in the direction of reforms. Since the 2000s, Nordic countries have experienced an increased impact of globalization and competition, mainly through international assessments that led to different levels of recentralization and a relative decrease in teacher autonomy.

Finland is an example with the highest level of teacher autonomy. It is also known for the strong cultural goal of schooling, equality, and having everybody integrated into the Finish society (Salonen-Hakomaki et al., 2016) and historically valuing and believing in teacher autonomy (Paradis et al., 2018; Sahlberg, 2011). Although Finland is not entirely immune to global trends of neoliberalism, unlike other European countries, such as Germany, it has placed a strong emphasis on school autonomy and has empowered teachers through curriculum development (Erss et al., 2016; Etelapelto et al., 2015;

Haapaniemi et al., 2021). Schools get the frame document and are responsible for developing curricula (Carlgren & Klette, 2008). Teachers are entirely free to determine the content, as well as assessment methods for their classrooms. In addition, accountability has no place in the educational policy discourse. National assessments are administered; however, the purpose differs greatly from those in the US, England, etc. They are used as a diagnostic tool to improve learning and have no implications for individual teachers.

Similar to Finland, the curriculum documents in Sweden give a substantial amount of freedom to teachers and see them as curriculum makers: the state develops the goals, and schools and teachers have to concretize them (Carlgren & Klette, 2008). However, unlike Finnish teachers, their autonomy has been constrained by stronger and increasing external demands and increasing responsibility for the outcomes of learning (Paulsrud & Wermke, 2020).

Teacher autonomy-related situation is quite different in Norway. Norway has historically been most accustomed to top-down and prescriptive curricula among other Nordic states (Bakken, 2019; Carlgren & Klette, 2008). Unlike Swedish and Finnish teachers, who are seen as curriculum makers, Norwegian teachers are constructed as curriculum deliverers (Bakken, 2019, Carlgren & Klette, 2008). In addition, Norway has recently placed a higher emphasis on teacher accountability and student outcomes. Unlike Finnish teachers, who are widely respected and trusted, Norwegian teachers have to a certain extent, lost high status due to their unfavorable performance in international tests (Mausethagen & Mølsted, 2015).

Teacher autonomy in post-Soviet countries has been strongly affected by the nature of the standard reform packages and by the consequences of the break-up of the Soviet Union and so-called teacher crises: drastic decline in teacher salaries, teacher shortages, the decreasing status of teaching professionals, lack of professional commitment (Silova & Kazimzade, 2010; Teleshaliyev, 2013; Steiner-Khamsi et al., 2008). In addition to these challenges to teacher professionalism, such legacies as the culture of compliance top-down management style, bureaucratic and central control with overwhelming paperwork, rigid requirements, the prevalence of collectivist thinking as opposed to individualism and practices of uncritical conformity to the decisions made by others have had negative consequences for teacher autonomy and agency (Ayubayeva, 2018; Khachatryan et al., 2013; Teleshaliyev, 2013). Despite this shared past, local educational stakeholders contested the standard reform packages as having different implications for teacher autonomy and agency. Responses to the proposed or imposed reform packages varied by socio-political and economic contexts (aspirations to join the EU, relations with Russia, presence and interests of non-Western players, the geopolitical situation, economic potential, etc.) (Silova et al., 2021). Overall, these legacies and contextual factors caused teachers' reluctance to willingly and self-confidently engage in the instructional process and use the freedoms (varying in their degrees across the countries) provided within the outcome-based centralized curricula.

Socio-cultural factors have also significantly determined the state of teachers in Arab-speaking countries. Participation in international assessments has created the desire among these countries to improve the achievement of their students by borrowing Western education decentralization reforms, including the creation of more than



previously available spaces of freedom for classroom teachers. However, the educational policies remain very centralized, including a highly prescriptive national curriculum. In addition, deeply embedded socio-cultural factors widely determine the absence of any agency among Arab educators and their strong preference to stay as passive recipients of educational reforms and be faithful implementers of top-down policies and regulations (Chaaban et al., 2021; Mustafa & Cullingford, 2008).

Political, social and cultural, and economic factors have greatly influenced shaping the vision of the role of teachers in South and East Asia countries. Education is seen as a driver of their economic development. As a result, instead of promoting activism or professional autonomy for educators, the government expects teachers to diligently follow the nationally outlined goals and dedicate themselves to the nation's particular needs and aspirations, including economic development (Ro, 2020). Even when the small space of freedom emerges, teachers are using it with the consideration of what needs to be done to meet the goal of centralized examinations (You & Morris, 2016).

### **Existing Conceptualizations of Teacher Autonomy**

Autonomy is one of the key aspects of the teaching profession. Scholars and practitioners started to actively discuss the idea of teacher autonomy as a significant point of focus rather than a supporting concept of learner autonomy in the 1990s (Benson, 2011; Ramos, 2006). In response to global and national educational reforms, teacher education scholarship gained momentum in the early 2000s (Salokangas & Wermke, 2020), along with numerous studies about teacher autonomy, conceptualization, and definitions used to vary a lot. Below I discuss two main conceptual approaches to teacher

autonomy: autonomy as a decision-making space and autonomy as a self-governance capacity. In addition, this section will also address the multidimensional nature of the concept of autonomy.

### **Teacher Autonomy as a Decision-making Space**

A review of the research on teacher autonomy, as well as conceptual differences drawn by various authors (Benson, 2016; Mausethagen & Molstad, 2015; Vangrieken et al., 2017), allows us to identify two major conceptualizations of this key term: (1) Teacher autonomy as a decision-making space, and (2) Teacher autonomy as a self-governance capacity. The former is a more traditional conceptualization and is informed by organizational theorists. They see autonomy as a key determinant of organizational efficiency and define it as a decision-making power and freedom to think and act (Luthans, 1992, seen in Friedman, 1999). As I will demonstrate below, educational researchers have extensively adopted this conceptualization for examining teacher autonomy. For example, Pearson and Hall (1993) define autonomy as teachers' perception about how much control they have of their work environment. Similarly, autonomy is seen as individuals controlling the terms and content of their work and other related issues (Molander & Terum, 2008, seen in Mausethagen & Molstad, 2015). Simpson et al. (2018) view autonomy as the systematic permission teachers are given within their educational systems. For Wermke (2013), autonomy is the scope of action that the state defines for teachers by its regulatory documents. Likewise, autonomy is seen as the discretionary space that reform provides (Ronnberg, 2007) or the power over the key decisions that affect various aspects of teachers' work (Ingersoll, 1996).

For a more nuanced understanding of autonomy as a decision-making space, it is important to acknowledge its two components: freedom of action (i.e., liberty to act) and freedom from control/constraints. Friedman (1999) analyzed the conceptualization of autonomy adopted by Charters (1976), one of the first researchers examining this construct. As Friedman (1999) argues, Charter's understanding of teacher autonomy is not about what teachers can do with the liberties granted to them. Instead, he conceptualizes teacher autonomy "as a shield against external pressures such as distrust, strong influence, control, excessive organizational demands, and pedagogical limitations" (p. 59). Another recent study (Tahirsylaj, 2019) takes a similar approach and defines autonomy as freedom from others' control over professional action or development. Priestley et al. (2015a) also emphasize the importance of freedom from control and define autonomy as a comparative absence of regulation.

In the continuing effort to define autonomy, one may consider the group of researchers that explicitly base their definitions on the premise of freedom of action. For example, Friedman (1999) conceptualizes autonomy as the power of teachers. He differentiates different degrees of autonomy and argues that a teacher with complete autonomy is "granted complete freedom to initiate and implement new ideas, programs, or curriculum within commonly accepted moral and legal principles" (p. 63). Conversely, teachers are believed to have no autonomy if they "are not authorized to take the initiative and are not given discretion in introducing changes in teaching methods and curriculum or any other elements of school life" (p. 64).

More recent studies also maintain the same line of definitions. Mills (2016) views autonomy as having the freedom to fulfill roles and responsibilities. Lundström (2015)

and Gurganious (2017) define autonomy as the amount of freedom and power teachers have to decide what teaching methods and content they use in their classrooms.

Frostenson (2015) and Gwaltney (2012) offer more detailed and elaborated definitions along the same lines of freedom of action. For Frostenson (2015), autonomy is "the freedom of professional actors to define the nature of professional work with regard to its formal contents, quality criteria, entry barriers, formal education, control mechanisms, ethics, et cetera." (p. 20). Quite similarly, Gwaltney (2012) defines teacher autonomy: "the degree to which teaching provides substantial freedom, independence, power, and discretion to participate in scheduling, selecting, and executing administrative, instructional, and socialization and sorting activities both in the classroom and in the school organization at large" (p. 22).

Erss and Kalmus (2018) intentionally adopt both liberty to act and freedom from constraints as constituting elements of their definition of autonomy. They argue that since freedom and control are in a dialectic relationship, researchers must thoroughly unpack both ends of the continuum to understand this complex phenomenon. To show this complexity, they explain how teachers may be given complete freedom to plan their teaching as they see it appropriate. Simultaneously, they may be deprived of the ability to make the best use of their creativity due to constraints, including centralized control, standardized testing, or time pressure to meet the curriculum goals in a predetermined period.

### **Teacher Autonomy as a Self-governance Capacity**

Research suggests that autonomy as a decision-making space, as discussed above, is only one component of teacher autonomy. It also includes the internal capability of an

individual to use and exercise the granted freedom productively (Benson, 2016). For example, Mausethagen and Molstad (2015) see this aspect of autonomy as "one's capacity to develop, safeguard, and justify one's knowledge base" and take responsibility for the performance (p. 31). Likewise, an autonomous person is defined as an individual who can reflect on their freedoms, engage in self-directed learning, and commit to exploring, changing, and growing (Dworking, 2015; Ramos, 2006). Bucelli (2017) differentiates between de jure and defacto autonomy to understand the two different aspects of autonomy. According to her, de jure autonomy is a set of rights granted to an individual for self-governance. On the other hand, de facto autonomy is understood as personal efficacy, i.e., the condition of a person having both competence and opportunity to use those rights productively.

Similar to Bucelli's (2017) research, a growing body of studies adopts a more comprehensive conceptualization of autonomy and argues that it conflates both components - freedom of action and capacity for action (Lennert da Silva & Mølstad, 2020; Ronnberg, 2007; Smith & Erdogan, 2008; Zeng, 2013). Ronnberg (2007) sees autonomy as a "function of both increased scope, as well as increased ability for action to increase the level of autonomy" (p. 218). Along the same lines, Helgoy and Homme (2007) criticize the narrow and static conceptualization of autonomy as granting teachers certain sets of rights does not mean that they will automatically use them for the best interests of their students. Moreover, the absence of control may result in some teachers "using the habitual forms of practice without much critical reflection" (Erss, 2018, p. 240). Mausethagen & Molstad (2015) also put forward the idea of a multidimensional conceptualization of autonomy with the following three components: (a) pedagogical

freedom and absence of control, (b) the will and capacity to justify the practices, (c) local responsibility. As they argue, these different components of autonomy are intertwined and need to be equally considered for a comprehensive understanding of teacher autonomy.

Based on a rigorous critical review of literature on teacher autonomy, Wilches (2007) argues for the necessity of separating the concepts of teacher autonomy, teacher competence, and motivation, with the latter two being more associated with the capacity-focused conceptualization of autonomy. Like Helgoy and Homme (2007) and Erss (2018), they also believe that the freedom to exert control is not always translated into exercising this freedom, i.e., autonomy, as teachers may lack the capacity, willingness, and motivation to do so. Conversely, teachers may lack the professional ability to use the freedom granted to them but have a strong motivation to exercise autonomy to develop competence. However, instead of incorporating all concepts into the construct of autonomy and generating confusion, Wilches (2007) suggests drawing lines of separation between autonomy, competence, and motivation. They define autonomy "either as a personal sense of freedom for professional action or as the power to exercise control in different school matters" (p. 268). Competence, motivation, and other personal capacity-related constructs can be regarded as factors that enhance or constrain professional action.

Finally, as Wilche's (2007) conceptualization demonstrates, autonomy can be defined as perceived or actual power to exercise control over teaching. For example, Short and Rinehart (1992) identify autonomy and decision-making power as two separate dimensions of the six-dimension teacher empowerment framework. They define decision-

making power as a level of teachers' engagement in decision-making. At the same time, autonomy is understood as a teacher's belief about the degree of control over different aspects of their work. Other authors also emphasize the importance of the "perception" component of autonomy (Friedman, 1999; Paradis et al., 2015; Parker, 2015; Pearson & Hall, 1993). Knowing what scope of action or what freedoms teachers are granted is essential. However, it is more critical to understand how teachers view the degree of this freedom and scope of actions, as their perceptions significantly influence their attitudes and commitment towards teaching and their self-confidence (Paradis et al., 2015; Pearson & Moomaw, 2005; Salokangas & Wermke, 2020).

### **Multidimensional Nature of Teacher Autonomy**

Cribb and Gewirtz (2007) developed a helpful conceptual framework that helps understand and examine the multidimensional nature of teacher autonomy and its relationship with control. Their framework is based on the relational understanding of autonomy. It implies that teachers' work is socially embedded, and their autonomy is always constrained to a certain extent. The authors differentiated between three major dimensions: (a) loci and mode of autonomy; (b) domains of autonomy, (c) loci and modes of control. As they argue, these three dimensions are complex elements and interact differently, leading to the multifaceted nature of teacher autonomy.

Multidimensional approaches proposed by other researchers (Frostenson, 2015; Wermke & Forsberg, 2017; Helgoy & Homme, 2007; Ingersoll, 1996; Friedman, 1999; LaCoe, 2006) are also helpful for enriching Cribb and Gewirtz's (2007) three-dimensional framework.

### ***Loci and Modes of Autonomy: Whose Autonomy?***

The first dimension considers who the agents are and how they exercise autonomy. Teachers are not the only agents in the educational context who hold or are expected to have autonomy. Cribb and Gewirtz (2007) also consider parents, students, and school management representatives or local government as important agents as individual, collective, or institutional agents. As they explain, teachers are collective agents when represented as teams within the schools or as members of teacher unions. Institutional agents include schools and their constituting bodies, local government agencies, and others. Another critical question to remember is how these different agents exercise their autonomy: individually, collectively, or on behalf of an institution. These distinctions are important as, in the context of educational reform, the autonomy of one agent exercised individually, collectively, or institutionally may have to strengthen or weaken implications for another agent's autonomy.

To understand the complexities of the debate on teacher de-professionalization, Frostenson (2015) also puts forward a three-tiered definition of teacher professional autonomy, which closely aligns with Cribb and Gewirtz's (2007) categories under loci and mode of autonomy. First, general professional autonomy refers to teachers' autonomy as the professional group and indicates their freedom to organize the framing of a teacher's work. This may include mandates related to the organization of the school system, curriculum, types and forms of control, organization of teacher education and entry requirements for teachers, etc. Second, collegial professional autonomy refers to the joint mandate of teachers as a school unit to decide and organize their professional work at the local level. Frostenson (2015) differentiates between two different sources of collegial autonomy: (1) delegation of responsibilities by the school management and (2)



collegial actions and decisions initiated and carried out by individual teachers. Third, individual autonomy is understood as an individual teacher's degree of freedom to influence her work at the classroom level. "This includes a choice of teaching materials, pedagogy, the mandate to decide on the temporal and spatial conditions of work, and to influence the evaluation systems of professional teaching practice" (p. 24).

Other authors propose a two-dimensional understanding of the concept regarding loci of autonomy. Wermke and Forsberg (2017) differentiate between institutional and service autonomy. The institutional dimension corresponds with Cribb and Gewirtz's (2007) institutional and Frostenson's (2015) general professional autonomy. Service autonomy refers more to the profession's practical dimension and conflates individual and collective/collegial autonomy from the abovementioned categorizations. Researchers also see autonomy as a continuum with two extremes of extended and restricted autonomy (assuming that autonomy always exists under certain conditions and complete freedom does not exist in reality). Depending on time and space, teachers may have constrained service and extended institutional autonomy, or vice versa (Wermke & Forsberg, 2017; Wermke & Hostfalt, 2014). They may also have both types of autonomy restricted, which would mean the full de-professionalization of teachers. The combination of extended service and institutional autonomy is another extreme and not plausible in the system of mass schooling "because teachers would then somehow be freelancers: they would not be related either to the control of the profession itself or related to any organization, in effect making individual contracts with their pupils. This is most likely the case if there is no systematized mass schooling system" (Wermke & Hostfalt, 2014, p. 69).

Helgoy and Homme's (2007) two-dimensional conception of autonomy includes individual and collective autonomy, claiming that depending on contextual factors, teachers may exercise a stronger collective or individual autonomy. In this case, collective autonomy also has aspects of Frostenson's (2015) general autonomy. Ingersoll's (1996) classification of individual and collective autonomy is more straightforward. The former indicates teachers' power teachers exercise individually for planning and teaching in the classroom, and the latter refers to teachers' collective influence over school-level policies.

Distinguishing between different types of autonomy regarding loci and mode is important for several reasons. First, it allows us to overcome the one-sided emphasis on individual notions of autonomy and recognize the differential impact of educational policies and interventions on different agents (Hermansen, 2017). Second, it helps understand how different agents exercise their autonomy and the implications of these diverse dimensions of autonomy for the educational system (Cribb & Gewirtz, 2007). Finally, this is a valuable tool for examining the relational and complex nature of the concept thoroughly and understanding how one group or individual's autonomy may affect others (Frostenson, 2015).

### ***Domains of Autonomy: Autonomy/Control over What?***

Cribb and Gewirtz (2007) highlight the importance of the domains, i.e., matters over which teachers and other agents may be autonomous or controlled. According to Ingersoll (1996), the answer to "autonomy and control over what?" depends on how different public education systems and researchers see teachers' roles in schools' core productive activities. For some, teachers' function is limited to the classroom zone, which

is about teaching and educational activities. Some authors also recognize the so-called school zone, which is about school management, planning, coordination, and resource allocation.

As Salokangas and Wermke (2020) claim, a substantial part of empirical studies on teacher autonomy mainly emphasizes the salience of the pedagogical practice at the classroom level, including teaching, planning, assessment, and curricula work. For example, Pearson and Hall (1993) see teacher autonomy primarily through the instructional component and identify the following domains across general teaching and curriculum autonomy: (a) selection of activities and materials, (b) classroom standards of conduct, (c) instructional planning and sequencing, and (d) personal on-the-job decision making. Vangrieken et al. (2017) propose a similar conceptualization. They make a distinction between content and pedagogical aspects of classroom practice and suggest ten domains: (1) Content aspect – preparing lessons, choosing topics and skills to be taught, decisions about the curriculum, choice, and use of textbooks, student assignments, setting goals for students, and student evaluation; (2) The pedagogical aspect – decisions about teaching methods and strategies, the use of time and planning, and managing student behavior. LaCoe (2006) and O'Hara (2006) stay with the same pedagogy-focused understanding of autonomy and build their concepts of autonomy in the following areas: curriculum, pedagogy, assessment, student discipline, classroom environment, and professional development. The last component of professional development is an addition compared to the domains developed by other authors I discussed above.

Other authors drew attention to the importance of non-academic aspects of teacher autonomy. Friedman (1999) and Wilches (2007) suggest the following four

domains: (a) Teaching and assessment (decisions about educational goals, content, methods, and materials, decisions about student evaluation, student discipline, and behavior, the physical environment in the classroom); (b) Curriculum development (teachers' interpretation and implementation of the formal curriculum); (c) Professional development (decisions about the subjects, content, time-schedules, and procedures of in-service teacher professional development and own engagement); (d) school functioning (decisions about school goals and vision, budget allocations, and school policies on classroom composition, student admission, etc.). Besides, Friedman (1999) differentiates between principle and routine decisions and suggests using the following types of decisions for examining teacher autonomy: principle pedagogical, principle organizational, routine pedagogical, and routine organizational. By the principal decisions, Friedman means decisions that affect fundamental aspects of teachers' work. In contrast, routine decisions refer to the ones that "are not intended to alter the organization's basic rules in any way" (p. 62).

Gwaltney (2012) proposes another expanded conceptualization of autonomy across different domains and levels. As their four factors of autonomy suggest, they incorporate both instructional and organizational aspects of teacher autonomy at both classroom and school levels: (1) classroom control over student teaching and assessment; (2) schoolwide influence over organizational and staff development; (3) classroom control over curriculum development, (4) schoolwide influence over school mode of operation.

Finally, in one of the most recent studies, Paulsrud and Wermke (2020) propose an analytical device for measuring, analyzing, and comparing teacher autonomy. Their

matrix incorporates the conceptualizations and suggestions of different scholars and puts forward four domains of autonomy: (1) educational domain: planning, instruction, and assessment; (2) social domain: socialization of students; (3) developmental domain: teachers' professional development; (4) administrative domain: use of resources, timetabling, salaries, etc. Besides, Paulsrud and Wermke (2020) distribute these domains across two levels: (1) classroom level (individual teacher's scope of activities within the classroom; (b) school level (scope of action of teachers as a collective). For example, as the authors explain, sanctioning a student is a social domain of the teacher's autonomy at the classroom level. In contrast, developing a student code is a school-level activity within the same dimension. Thus, this matrix addresses both the second and the first dimensions identified by Cribb and Gewirtz (2007) – loci and mode and domains of autonomy.

Such a conceptual differentiation between the domains of autonomy is very important as it helps capture the complexity and variations of teacher autonomy within this dimension. Teachers may be targets of tight prescriptions in one area while enjoying broader autonomy in other directions (Cribb & Gewirtz, 2007). It also clarifies conceptual confusion and unpacks how different authors define and measure autonomy in their empirical studies (Paulsrud & Wermke, 2020).

### ***Loci and Modes of Control***

The third dimension in Cribb and Gewirtz's (2007) framework refers to the loci and modes of control. It is aimed to answer the following critical question for the research on teacher autonomy – who are the agents that have the power to exercise control, and how do they exercise this control? Regarding teachers, agents of control can

be representatives of school leadership, local or central government, parents, teacher unions, etc. Sometimes, the loci of control are not explicitly known. For example, it may not be exercised purposefully and in an organized manner by a specific organization or a group of individuals. The authors of the framework also explain how different agents may exercise their control by using different mechanisms of influence. Also, the style and degree of control may vary from bluntly coercive to gentle persuasive.

Hudson (2007) describes different control mechanisms exercised by the central government: (a) curriculum control, (b) school evaluations and inspections, (c) standards, assessments, and accountability, (d) self-regulation. The degrees of control may vary. According to her, the degree of curriculum control depends on the nature of the curriculum. Control is more subtle if it is a regulatory framework and leaves much room for interpretation. On the other hand, more prescriptive curricula are examples of more assertive and more direct control. The government can also exercise indirect control by providing different guidance documents to schools on interpreting and implementing the curriculum. To differentiate the degrees of control, Hudson (2007) also brings examples of systematic evaluation practices in the form of external evaluations and self-evaluation. Depending on the goals and specific procedures of these evaluations in various countries, they may be more subtle (e.g., in Finland) or very direct and associated with severe consequences for improper performance (e.g., in England). Hudson (2007) brings up another interesting point about how output control of schools and teachers through standardized assessments generates new control mechanisms for other agents in the system, such as parents and the public in general.

Studies on teacher autonomy differentiate between input control, which targets resources for instruction, and output control, focused on performance measurement (Helgoy & Homme, 2007; Hopmann, 2003; Wermke & Forsberg, 2017). Hopmann (2003) uses concepts of product and process controls of the curriculum. The former is an input-based control, also known as Didaktik or European approach (Wermke & Höstfält, 2014), and uses plans and frameworks as the primary control instrument. If the procedures taken by the schools and teachers are aligned with these plans and frameworks, then different results in student achievement are allowed. The latter is an output-based control, known as the Anglo-American curriculum approach (Wermke & Höstfält, 2014), and uses standards, outputs, and achievement results as the primary control instruments. In contrast to the input-based approach, it allows differences in the process and procedures as long as student achievement results meet the predetermined expectations.

Types of control (input vs. output) have implications for teachers' autonomy. As Hopmann (2003) described, the Didaktik system allows for a considerable amount of teacher autonomy as it has weak control of the educational process and almost no external control over the outcomes. Wermke & Höstfält (2014) differentiate these implications by types of autonomy – institutional vs. service autonomy. They claim that in process-related curriculum evaluation, teachers have extended institutional autonomy and restricted service autonomy. In contrast, in product-oriented curriculum evaluation systems, teachers have a combination of extended service autonomy (as they have the right to broad choice content and methods) and decreased institutional autonomy (as their products are evaluated by standardized tests).

Such a multidimensional conceptualization of autonomy and control has implications for policy analysis and cross-national evaluations regarding the consequences of new forms of regulations. As the overview of three dimensions suggests, instead of presenting autonomy and control-related inquiries as a black-and-white issue, we need to ask the following complex question: "What effects do different combinations of forms and styles of regulation each has on the autonomy of the different individual, institutional and collective agents, in relation to each of the different relevant domains?" (Cribb & Gewirtz, 2007, p. 210).

### **Measuring Teacher Autonomy**

Several empirical studies have focused on developing and validating the research instrument for measuring teacher autonomy. These studies propose different tools and scales for the construct of teacher autonomy, in line with their conceptualization of this important concept under examination. The Sense of Autonomy Scale (SAS), developed by Charters (1976, seen in Friedman, 1999), focused on the personal sense of freedom from external interference. It included the following elements: (a) Freedom to select work techniques, (b) Freedom from distrust by administrators or colleagues, (c) Freedom from an administrator or colleague influence, (d) Freedom to control the pace of student Work, (e) Freedom from excessive school level organization of instruction, and (f) Freedom in relationships with students (seen in Friedman, 1999).

Another instrument, the Teaching Autonomy Scale (TAS) by Pearson and Hall (1993), measured teachers' perceptions about whether they can control their work environment. For example, one of the items of the 35-item questionnaire reads as follows: "What I teach in my class is determined for the most part by myself" (Pearson &



Hall, 1993, p. 174). The instrument identified two dimensions of teacher autonomy – general autonomy indicating "issues concerning classroom standards of conduct and personal on-the-job discretion" (p. 177), and curricular autonomy, referring to "issues concerning the selection of activities and materials and instructional planning and sequencing" (p. 177) autonomy than their colleagues at elementary and high schools.

Instead of focusing on freedom from external constraints, Friedman's (1999) Teacher Work-Autonomy Scale (TWA) measured autonomy as a decision-making power and any kind of organizational behavior. For example, one of the questionnaire items asked teachers to reflect on the following statement: "Teachers formulate and try out innovative curricula" (Friedman, 1999, p. 72). It also took a broader approach and measured teacher autonomy in the following four domains: curriculum development, teaching and assessment, professional development, and school functioning (same as administrative matters).

As some of the authors mentioned above, LaCoe (2006) operationalized autonomy as "the amount of latitude or control a teacher has over the following areas: curriculum, pedagogy, assessment, classroom environment, and professional development" (pp. 41-42). However, LaCoe examined not only the autonomy teachers perceived to have but also the amount of autonomy they desired. Similar to Friedman's (1999) approach, items under the autonomy *have scale* were formulated as organizational behavior: "I determine the consequences of negative student behavior that occurs in my classroom" (LaCoe, 2006, p. 62). In contrast, the corresponding item under the autonomy desire scale read, "Teachers should decide how to punish any student discipline infractions that occur in their classrooms" (p. 62). According to LaCoe (2006), it is

interesting to measure actual versus desired autonomy during the accountability area and see what amount of autonomy the teachers of tested subjects' would like to have.

Paulsrud and Wermke (2020) have developed one of the most recent instruments for measuring, analyzing, and comparing teachers' perceived autonomy. Their questionnaire aimed to understand teachers' perceptions about who makes important decisions at different levels in different domains. They proposed the following four domains of autonomy: educational (lesson planning, delivery, and evaluation), social (discipline policies, tracking, special needs), developmental (formal professional development), and administrative (scheduling, time, and financial resources). They saw teacher autonomy being exercised at two levels: individually in the classroom or collectively in school. Also, they identified four types of potential decision-makers: individual teachers, teachers collectively, school management, and actors outside the school. Such an instrument allowed for a more nuanced measurement of teacher autonomy by different domains and levels.

An overview of the conceptualization of autonomy and its measures shows that autonomy is not a single trait but a multifaceted concept with different domains, levels, and dimensions. Therefore, it is challenging to create an instrument that captures the complexity of this concept. As mentioned above, as a decision-making capacity, autonomy can be viewed as the freedom to act, control of tasks, or freedom from pressure and constraints. Existing measures favor one of these conceptualizations rather than encompass them comprehensively. For example, Friedman (1999) conceptualizes autonomy as the freedom to or power to act, while for LaCoe (2006) and Pearson and Hall (1993), autonomy is more about control of tasks, and Charters (1976) focuses on

freedom from pressures. The same applies to capturing multiple domains and zones of autonomy. As the overview of the instruments also suggests, some measures take a more comprehensive approach (for example, Friedman, 1999; Paulsrud & Wermke, 2020) than others (such as Charters, 1976).

### **Autonomy as a Source of Professionalism**

Autonomy is often discussed as a necessary condition for professionalism in any field. Dewey (1903) sees it critical for teachers to have continuous opportunities to affect the school system by their judgment and, instead of being recipients of prescriptions, engage in intellectual discussions and initiatives. Dewey (1903) also points out that:

The dictation, in theory at least, of the subject-matter to be taught, to the teacher who is to engage in the actual work of instruction, and frequently, under the name of close supervision, the attempt to determine the methods which are to be used in teaching, mean nothing more or less than the deliberate restriction of intelligence, the imprisoning of the spirit. (p. 196)

Gorman and Sandefur (2011) identify four essential characteristics of true professionalism: expert knowledge, the use of professional discretion and autonomy, normative orientation and community, and (4) status, income, and rewards. Similarly, autonomy is considered necessary for teaching and professionalism in education (Cribb & Gewirtz, 2007). Scholars also see it as a central factor that draws a line between the understanding of teachers as technicians and as professionals. According to Webb (2002), those who approach teachers as technicians think all teaching requires is knowledge of the subject matter and basic skills. From this perspective, teachers must adhere to a set of guidelines and use teacher-proof curricula. Opposite to this, teachers as professionals are

expected not only to possess talent and skill but also "have the autonomy to make decisions that marry skills with knowledge" (Webb, 2002, p. 50). In contrast to teacher technicians, teachers as professionals engage in complex thinking "make normative judgments about the kinds of attributes, dispositions, and values that they want to inculcate in their students and the kinds of formal and informal curricula, pedagogies and forms of assessment that are worthwhile "(Cribb & Gewirtz, 2007, p. 206). Along the same lines, Berliner and Glass (2014) argue that there is no use in teaching teachers how to promote higher-order thinking skills unless they have the freedom to "have the autonomy to create classroom environments that encourage exploration, dialogue, and conversation" (p. 90).

Likewise, Biesta (2015a, 2015b) and Freidson (2001) see professional judgment and autonomy as necessary conditions for teaching. Biesta (2015a) explains the multidimensional purpose of education to demonstrate the centrality of judgment in teaching. According to him, education has three domains of purpose: qualification, socialization, and subjectification. This means that teachers are expected not only to help students acquire knowledge, skills, and dispositions (as part of the qualification dimension) and to engage extensively with the student as a person. Despite the potential for peaceful synergy, these three domains also often come into conflict and require good teachers to make professional judgments. This means that teachers will have to decide when to make education "flexible, personalized, and tailored to individual students" for "promoting creative thinking or generative actions" and when to keep it "strict, structured and general" and act in a prescribed way (p. 80). Thus, good teaching goes beyond competence and technical knowledge. It requires professional wisdom and autonomy for

"judgment about what an educationally desirable course of action is in this concrete situation with these concrete students at this particular stage in their educational trajectory" (Biesta, 2015b, p. 5).

In their book, "Professional Capital: Transforming Teaching in Every School," Hargreaves and Fullan (2015) emphasize the importance of judgment for the teaching profession. However, they question the appropriateness of this judgment taking place at the individual level. The authors introduce a concept of professional capital (as opposed to "business capital" understanding of the teaching profession), which in addition to human capital (the talent of individuals) and decisional capital (the wisdom and competence to make sound judgments about learners) consists of a social capital (collaborative power of the group). According to them, the time of individual classroom autonomy with unquestionable knowledge and expertise of an individual teacher has gone. Instead, they emphasize the importance of collective autonomy and responsibility, allowing teachers to enrich and inform their professional judgments through the insight, knowledge, experience, and support of their peers (Hargreaves & Fullan, 2015).

### **Teacher Autonomy Enhancing Student Learning and Equity**

The existing body of literature is quite contradictory about the implications of teacher autonomy for student learning and achievement. On the one hand, researchers value teacher autonomy for its potential to positively improve learning experiences and development. Another body of literature shows concern about the possible risks associated with teachers' freedom and calls for caution in autonomy-supportive interventions.

Teacher autonomy-supportive literature refers to indirect and direct mechanisms that mediate this desirable relationship between autonomy and student learning and achievement. Among other rationales, scholars emphasize the power of autonomy to foster well-being and job satisfaction among teachers. As Cribb and Gewirtz (2007) argue, "for the individual teacher freedom to decide what to teach and how to teach it and freedom to play an active part in decision-making about the conditions in which s/he teaches can be seen as a vital source of job satisfaction and physical and mental well-being" (p. 206). This argument builds on a self-determination theory, a macro-theory on human motivation (Deci & Ryan, 2000), that identifies three universal psychological needs necessary for effective functioning: feelings of competence, autonomy, and relatedness. They facilitate intrinsic motivation, which is critical for individuals to maximize their potential. Intrinsically motivated people experience more confidence, interest, and excitement in what they do and self-endorsement of their activities (Deci & Ryan, 2008; Ryan & Deci, 2006; Vansteenkiste & Ryan, 2013). Conversely, feelings of lack of power among teachers over homework policy and other classroom instruction matters increase their frustration and anxiety (Mayer et al., 2013). Moreover, the perception of lost autonomy negatively affects their self-confidence and makes them question their self-worth as a professional (Paradis et al., 2015). Therefore, it is reasonable to assume that "what is good for teachers is good for learning" (Cribb & Gewirtz, 2007, p. 206). When educators feel autonomous and supported, they improve their efforts (Marks & Louis, 1997) and offer more adequate education to their students with increasingly diverse needs and backgrounds (Hyslop-Margison & Sears, 2010, Prichard & Moore, 2016).

In addition, autonomy positively influences the learning process in the classroom by reinforcing meaningful continuous professional learning among teachers. Teachers who can control decisions over their professional development activities and are free of others' judgments and directions can engage in deep and sustained learning and better tailor new strategies to their students' needs. Contrary to this, the ones lacking space for independent initiatives show less enthusiasm for ongoing professional development (Forsberg & Wermke, 2012) or are involved in professional learning "in a ritualistic fashion," gaining only superficial knowledge of the studied matter (Hargreaves et al., 2013).

Forsberg and Wermke (2012) make an interesting contribution to understanding the relevance of autonomy for teachers' continuous professional development. Investigation of knowledge sources of assessment among German and Swedish teachers revealed educators' preference for those sources where they could act more autonomously and make selections and learning experiences by using their discretion. These sources were cooperation with colleagues and various forms of self-study rather than centrally or regionally organized continuous professional development activities. Also, as Forsberg and Wermke (2012) demonstrate, the level of autonomy and space for professional discretion in assessment influenced teachers' motivation to engage in continuous professional development. German teachers with more autonomy were more active in their professional development efforts. On the other hand, Swedish teachers with "the greater dependence on the assessment decisions made outside the individual classroom" (Forsberg & Wermke, 2012, p. 755) showed less enthusiasm for continuous professional development.

The authors also establish a more direct link between teacher autonomy and effective classroom practices. According to LaCoe (2008), one of the major advantages of teacher autonomy is that it creates opportunities for a more individualized approach instead of using one size fits all strategy and a top-down delivery model. Schools with empowered teachers are better positioned to diagnose and address the specific needs of their culturally diverse students. The potential of autonomy for such context and person-specific judgments is particularly critical for students from disadvantaged families (Marks & Louis, 1997; Prichard & Moore, 2016; Salokangas & Ainscow, 2017; Skaalvik & Skaalvik, 2009). Such judgments allow teachers to target all types of resources associated with learning and teaching to students most in need (Cribb & Gewirtz, 2007). Conversely, such constraining factors to autonomy as testing, prescribed curricula, and pacing standards seriously limit teachers' ability to meet the students' diverse needs, especially those with disadvantaged backgrounds (Olivant, 2015).

A further advantage of strengthening and developing teacher autonomy is its potential for enhancing and developing student autonomy (Benson, 2011; Leroy et al., 2007). It is believed that only teachers who themselves enjoy a reasonable degree of autonomy can serve as role models for students and raise them as autonomous and reflective citizens (Hyslop-Margison & Sears, 2010; Ramos, 2006). Benson (2011), a big advocate of learner autonomy, also sees teacher autonomy (both the freedom and internal capacity) as a necessary condition for developing learner autonomy. Similarly, Lawson (2004) emphasizes how only empowered teachers can pass control over learning to the students and turn it into a collaborative effort rather than forcefully imposing the knowledge. Conversely, Pelletier and Sharp (2009) explain how the reduced perception



of autonomy among teachers (caused by various administrative pressures) affects teacher behaviors in the classroom and students' motivation correspondingly. As teachers feel a lack of autonomy, they eventually start using more controlling behaviors toward their students and become less autonomy supportive. This makes a big difference for the students in the classrooms. Instead of being more responsive, supportive, and explicative and allowing students options to choose from, teachers take complete charge of the teaching and learning process, lead students to the correct answers and use pressure for motivational strategies (Assor et al., 2005, seen in Pelletier and Sharp, 2009). These relations are reciprocal too. As teachers controlling behavior negatively affect students' behavior motivation, the latter has negative consequences for teacher behaviors. A study of Chinese teachers and students indicated the same relationship. Students of teachers who had little freedom to make instructional decisions and were unused to learner-centered methodologies tended to have difficulties in taking responsibility for their learning and mostly reacting to external motivation mechanisms (Sinclair, 2009).

Teacher autonomy also has the potential to positively affect student learning by improving achievement standards (Ingersoll, 1997; Porter, 1989). Teachers can make the best-informed decisions for those they teach (Ingersoll, 1997). Teacher engagement can mitigate two major risks associated with externally imposed standards. First, external standards may contain inappropriate content for all students or some groups of students. Second, external standards may decrease intrinsic motivation among students and teachers as they may have to learn and teach the content they believe is inappropriate (Porter, 1989). Therefore, teacher participation in the development of standards can make

them more applicable and relevant for their students, automatically leading to strong intrinsic motivation and improved learning for students.

In addition, autonomy can play an essential role in improving student learning by allowing those directly engaged in teaching to be more creative and experiment in search of the best possible approaches for the benefit of their students (Cribb & Gewirtz, 2017). When teachers feel a lot of pressure, they have little or no enthusiasm to demonstrate originality in their work and efforts (Moomaw, 2005). Removal of the fears associated with the performativity culture and allowing educators to teach creatively can significantly foster creativity in the classroom and "deepen students' learning, broaden their higher-level thinking skills, and enhance their enjoyment of learning and motivation to learn" (Olivant, 2015, p. 128).

### **Teacher Autonomy Creating Risks for Student Learning and Equity**

In addition to the benefits teacher autonomy can bring to the students in the classrooms and equitable educational opportunities, on the negative side of the spectrum, the body of research raises red flags about the threats teacher freedom can create directly or indirectly. Not all researchers believe autonomy is a universal need and always desired condition, as conceptualized by Deci and Ryan's (2000) self-determination theory. Some teachers simply do not want to have the autonomy they are given. Instead of making decisions on critical issues, they prefer to be guided by their headmasters or inspectors (Moomaw, 2005). In fact, Hulpia et al. (2009) found that teachers' organizational commitment was more influenced by how teachers perceived the support they received from the school leadership team than by how they viewed empowerment in school decision-making.

Even if autonomy is considered a universal need and equally desired by teachers of different mastery levels, it may affect the system differently, given teachers' different types and personalities. As Skaalvik and Skaalvik (2014) speculate, although studies demonstrate that decision latitude (i.e., autonomy) and job satisfaction are positively related for all teachers, this happens through different processes. Also, this relationship may look different in the long-term perspective as opposed to the short-term one. Teachers with high mastery levels may desire autonomy to freely use the best of their potential to teach and respond to their students' needs and experiment with new ideas and practices. Such opportunities and processes will lead to high engagement and commitment. On the other hand, educators with lower mastery level may desire autonomy for avoiding challenges and hiding self-perceived shortcomings. Thus, autonomy can become a self-protective strategy that may increase the level of job satisfaction among teachers. However, this may act as a barrier to learning and development in the long run, consequently negatively affecting job satisfaction and, most importantly, the teaching and learning process for their students.

Given the multifaceted nature of autonomy, it is important to understand how teachers perceive the significance of autonomy across different domains. Based on the analysis of previous studies, Moomaw (2005) argues that teachers' desire for autonomy differs by the area of discretion. For example, while most of them find autonomy in planning and instruction important and valuable, they are more hesitant and less appreciative of having autonomy and decision-making power on school-management matters. Strong and Yoshida (2014) indicated two reasons why instructional / classroom autonomy could be more valued and desired by teachers – they may feel more

professionally able for classroom-related autonomy, and they may find non-classroom-related matters less important for their professional activities. Ingersoll (1996) brings a different perspective and argues that increased curriculum autonomy solely may worsen teachers' relationships with students unless their decision-making power over socialization activities (i.e., discipline and sorting) increases in parallel. Thus, the effect of autonomy on student-teacher relationships, also affecting the overall learning process, may depend on the type of combination of the autonomy domains.

Other studies indicate various reasons due to which teachers' desire for autonomy may vary. For example, Chinese teachers showed no willingness to have curricula autonomy as these changes were inconsistent and unexpected, resulting from governmental changes. Their desire was also decreased by the partial flexibility given to them, as subject standards and content were strictly defined and did not give them much choice of actions (Hong & Youngs, 2016). Another study of Estonian, German, and Finnish teachers (Erss, 2018) revealed different attitudes toward autonomy in different countries. Specifically, in addition to the level of mastery and professionalism, the study unpacked such conditional factors as lack of support and excessive control mechanisms. For example, Finnish teachers were excited to have a lot of freedom in their curriculum and classroom decisions, which Erss (2018) attributes to their high level of professionalism. On the other hand, Estonian teachers expressed concerns about forced autonomy and did not value their freedom in planning and instruction. According to Erss (2018), this is due to the lack of resources and very tough control mechanisms and regulations of personal answerability for student achievement.

As Skaalvik and Skaalvik (2014) also recognize, autonomy is not always a universally desirable thing, as teachers get not only freedom but also responsibility for how they use their freedom. Accountability in the form of centralized exams made IB teachers think of autonomy as a double-edged sword (Kelly, 2019). Similarly, a study of elementary teachers in Cyprus (Phillipou et al., 2014) indicated that teachers preferred less freedom and autonomy due to practical and professionalism-related reasons. On a practical side, they wanted to avoid personal costs of autonomy - stress and heavy workload associated with preparing the curriculum materials. They were also concerned about the lack of school resources supporting new teaching methods. On a professional side, they preferred to have the centrally provided curriculum as a shield of protection against interference and doubts coming from the parents. Thus, due to such reasons as the level of professionalism, support systems, and control mechanisms, autonomy can be more a source of stress and anxiety than the well-being and job satisfaction, having implications for the teaching process and students they have in their classrooms.

Other authors express more serious concerns about the potential of autonomy to increase the learning process. Instead of improving their practices and thinking of innovative solutions, some teachers may maintain the status quo and follow the old patterns they are used to without any critical evaluation (Erss, 2018; Johnson, 2016). Teachers may be unable to take advantage of the freedom due to their lack of competence. For example, Ozturk (2012) examined Turkish teachers' classroom practices and found that despite autonomy (mainly resulting from structural looseness) in selecting content, materials, and teaching techniques for classrooms, not all teachers engaged in practices addressing student needs and their learning styles. This was particularly true for

the students with special needs. Ozturk (2012) attributes these differences in the effective use of flexibility and autonomy to different levels of teacher professionalism. As another example, when exploring the experiences of teachers from New Zealand, Hood (2020) arrived at an interesting conclusion: rather than relying on the standardized test results, a key component of the neoliberal educational reform, the government of New Zealand instituted Overall Teacher Judgements (OTJs) for measuring student success. However, analysis of the teacher practice showed that teachers did not have sufficient knowledge and expertise to apply OTJs. Due to various pressures, teachers may also use their autonomy for the wrong purpose. As the study by Hong and Youngs (2016) demonstrated, instead of enriching the curriculum, Chinese teachers simplified it as much as the content requirements of the entrance exams allowed.

Similarly, Woessmann (2005) sees the danger of opportunistic behavior to the proper use of autonomy. For him, the power of local and contextualized knowledge is a strong argument supporting autonomy. Locals are more knowledgeable as compared to the central ones on how to provide education services more efficiently. Correspondingly, teachers have good lead knowledge of delivering instruction most productively and how to respond to student needs best. "This local knowledge lead can make provision by a local agent more efficient than by a central planning authority" (Woessmann, 2005, p. 147). However, he also argues that local actors will only use this local lead knowledge to advance the quality of the final product if they have the incentive and desire to do so. According to him, the danger of opportunism mainly associated with divergent interests and asymmetric information shall always be overseen carefully, especially when the local actors see no real risk of their opportunistic behaviors being noticed and penalized.

A thorough examination of research also allows for identifying counterarguments to autonomy, having the potential to advance student learning through continuous and meaningful professional learning. Specifically, autonomy is believed to lead teachers to work alone, in isolationism (Anderson, 1987). Marks and Louis (1998) found that teachers who were empowered and had autonomy over classroom practices did not necessarily feel connected to their peers. "The classroom door provides a measure of autonomy for teachers, but it also fosters isolation, limits feedback and performance, and promotes staleness" (McLaughlin et al., 1986, p. 423). Teachers feel reluctant to give up their autonomy and work-alone mode, believing collaboration is difficult, time-consuming, and unproductive. It challenges individual freedom and questions their competence (Bakken, 2019; Bunker, 2008; O'Hara, 2006). Moreover, autonomy is often valued for removing the need for teachers to get together to discuss educational issues, saving planning time from unnecessary meetings, reducing coordination costs, and easing interpersonal tensions (Anderson, 1987). Also, O'Hara (2006) argues teachers who are afraid of self-reflection and critical evaluation by others may also purposefully use autonomy and isolation as a shield to protect their self-esteem from the perceived dangers of collaboration. Such isolation and practices of closed doors create a severe threat to cooperation and collaborative work (Gajda & Koliba, 2008). As a result, their learning and development process is more like a trial and error. How much they learn and grow and improve instructional practices in their classrooms depends solely on their personal abilities to independently grasp and master instruction-related content and skills O'Hara (2006).

Can autonomy be considered beneficial for classroom teaching if we keep understanding it as a degree of individual freedom and discretion? Hermensen (2017) argues that professional autonomy has been challenged differently recently. Policy documents (such as standards, curriculum, etc.), expectations of research-based practices, and systematic use of data are increasingly used as professional knowledge. This pattern has been exacerbated by intensive policy borrowing and digitalization. Individual autonomy is insufficient as adjusting the resources generated from different epistemic and methodological perspectives can be a very demanding and unrealistic task for teachers. Therefore, Hermansen (2017) believes that more collective efforts should be encouraged instead of engaging with these knowledge landscapes as an individual responsibility. Exercise of collective autonomy will be an effective strategy for putting teachers' voices in the classrooms by collectively defining what kind of resources and for what purposes should be used.

One of the arguments against teacher autonomy, especially from those who are committed to justice in education, is the risk of producing and perpetuating inequities in classrooms (Cribb & Gewirtz, 2007). In their study, Murphy et al. (1986) traced a pattern of lower-quality instruction provided to lower-ability groups as opposed to higher-ability ones. Based on the findings of various studies, including one by Murphy et al. (1986), LaCoe (2008) argues that autonomy does not necessarily lead to improved outcomes for all students as teachers may decide to use their autonomy negatively. For example, instead of taking advantage of the freedom to tailor instruction to their students' needs, teachers may deny access to rigorous curriculum to some students based on their perceptions of these students' abilities. Teachers can use their autonomy both to empower



and marginalize (LaCoe, 2006). Therefore, under the conditions of a high level of autonomy, there is a risk of lower-ability track students being deprived of a sufficient amount of opportunity to learn. Thus, the more decisions teachers can make about what to teach and how to teach, the higher the potential for inequity in the classrooms (LaCoe, 2008).

Based on Boote's (2006) description of the necessary competencies for teachers, teachers may end up misusing autonomy due to the lack of the ability to distinguish relevant idiosyncrasies of their students and base their decisions upon irrelevant ones, such as gender, ethnicity, socio-economic status, etc. Also, teachers may lack moral competencies and fail to make the right and ethical decisions about the students and their learning.

### **Autonomy – Suitable to All Teachers?**

Many researchers question the universal value of autonomy and invite consideration of various individual and context-related factors. These factors explain the risk of giving teachers the freedom to act that exceeds their actual capacity to act. As discussed above, teacher mastery and professional maturity are among these conditions (Anderson, 1978; Ashley, 2012; Erss, 2018; Grant et al., 2020; Littlewood, 1996; Strong & Yoshida, 2014; White, 1992). Even though younger teachers are more open to changes than their older colleagues, the lack of experience and expertise does not allow them to think critically and "surrender to imposed ideas and the obedient use of the method" (Tickle, 2000, p. 69). Also, inexperienced teachers are more concerned with survival issues. In contrast, those with higher competence have both self-confidence and the ability to act independently, purposefully, and creatively to address their students' needs.

Therefore, autonomy is more important for experienced teachers than inexperienced ones (Burden, 1981).

Boote (2006) proposes the most comprehensive list of competencies teachers need to have to be able to make appropriate curricula and instructional decisions: (1) ability to distinguish relevant idiosyncrasies of their students to avoid discrimination; (2) competence in subject matter; (3) pedagogical content knowledge to organize scope and sequence of instruction meaningfully; (4) general pedagogical skills for maintaining a productive learning environment; (5) moral competencies to make the right and best decisions for students; (6) competence in interpersonal communication to explain and defend their curricula decisions. Boote (2006) argues that in addition to these competencies, teachers should have self-control (to manage the fear or anxiety when using a new curriculum), as well as procedural independence (to make curricula decisions with no or little consultation with supervisors) and substantive independence (to critically evaluate and make the best decision by applying various criteria). Lack of any of these competencies may cause serious problems such as discriminating against students, wrong instructional and assessment methods and resources, and losing control of a productive learning environment (Boote, 2006).

This consideration can be supported by Fox's (2012) developmental perspective and framework of a "graduated autonomy" concept, which implies that autonomy is incremental, proportional, i.e., "scaled to life conditions" (p. 4) and is related to the life stage during which it occurs. Huberman (1989) puts the teaching profession in a similar developmental perspective. Throughout their career, teachers go through such stages as survival and discovery, stabilization, experimentation/activism, self-doubt, serenity,

conservatism, and disengagement. Teachers' desire for autonomy varies across these stages, achieving the highest level at the later stages of their professional development.

Grant et al. (2020) believe that the theory of development, as described above, closely corresponds with the developmental needs of teachers as they advance in their careers. Therefore, they argue that "mismatched levels of autonomy can lead teachers to believe that their professional growth, teaching practice, and ability to form connections with students is being undermined" (p. 104). They propose a Graduated Teacher Autonomy Framework that matches teacher proficiency levels across four teaching domains to the degrees of autonomy. This model makes autonomy not only case-specific (i.e., teacher-specific) but domain-specific as well. A teacher can be granted higher autonomy in the domain where they demonstrate a higher level of proficiency. They argue that such differentiated, incremental, and domain-specific approaches protect teachers from the freedom and scope of discretion they do not desire or cannot handle and maximize their professional effectiveness. Anderson (1987) takes a similar position and argues that the autonomy teachers are granted should mirror their experience, expertise, and status.

Boote (2006) supports the idea of graduated autonomy by proposing three levels of professional discretion ability teachers go through throughout their career. Procedural professional discretion is the lowest level of discretion that is expected from recent graduates of teacher pre-service programs and implies the ability to develop a minimally coherent curriculum to teach. The next level of discretion is substantive professional discretion indicating teachers' ability to evaluate and modify their instructional decisions. Finally, innovative professional discretion is believed to be a long-term goal of teacher

development as it encompasses "the ability to go beyond merely choosing among established and sanctioned curriculum options to creating new curricular–instructional practices that ameliorate the dilemmas of their domain of curriculum practice" (p. 468).

The idea of graduated autonomy can also be supported by ecological (Biesta & Tedder, 2006; Priestley et al., 2015a) and subject-centered socio-cultural (Etelapelto et al., 2013) frameworks of the agency. According to these frameworks, one of the critical conditions for achieving agency, i.e., exerting purposeful action, is learning from past experiences and utilizing professional knowledge and competencies. If teachers have no or limited professional past and a sufficient level of competence, they will not be able to use it as an enabling resource for exercising agency. Therefore, other factors, one of which could be more temporary support and guidance, can compensate for the lack of experience.

The findings of various studies support the need for graduated autonomy for the best benefits for teachers. For example, Gwaltney (2012) found that compared to non-tenured teachers, tenured educators with more experience perceived autonomy as more important to the motivating potential of teaching. Per Kauffman et al.'s (2002) findings, early-career teachers from Massachusetts did not appreciate freedom and opportunity for control of curriculum and instead asked for more guidance and prescriptions. Lambeth (2012) also found strong guided support with mentoring and other support systems critical for beginner teachers. Based on the findings of the study of music teachers in Finland, Juntunen (2017) also emphasized that the more freedom teachers have, the better knowledge and mastery they need.

Given the levels of competence and professionalism (along with other individual and contextual factors), teachers may position themselves differently for the same opportunity of granted freedom in curriculum and instruction. Phillipou et al.'s (2014) findings indicate that some teachers took a spectator position and requested prescriptions from experts. Another group, so-called receivers, had some desire for limited autonomy right to choose from ready-made instructional resources. Implementers saw themselves as the agents of change but more at the classroom level, willing to develop teaching materials in collaboration with colleagues with higher professionalism and expertise. Finally, those teachers who positioned themselves as reformers showed confidence in their knowledge and abilities and were willing to co-develop the curriculum in an equal collaborative framework.

Another study by McGrath (2006) helps us understand how teachers perceive textbooks and how they position themselves in relation to them. Using metaphors, teachers demonstrated four types of views of textbooks: guidance, support, resource, and 'constraint.' The guidance was associated with a very low level of autonomy, i.e., relying on the book too much. At the same time, the view of constraint expressed teachers' high level of autonomy and critical view. Although McGrath (2006) does not analyze the views teachers held by their level of professionalism, one can speculate that competence and mastery, along with other factors, greatly influenced how teachers viewed and used textbooks in their instruction.

As Ronnberg (2007) argues, autonomy is the function of both the freedom of action and capacity for action. Teacher competence and professional ability constitute one of the major factors of the capacity for action. Therefore, to increase the potential of

autonomy to benefit both the teachers and students as well as to avoid undesired consequences, the increase in the space of discretion should be accompanied by the corresponding rise in teacher competence and qualifications.

### **The Impact of Teacher Autonomy on Student Achievement and Achievement Gap – Existing Empirical Evidence**

The literature is not replete with studies directly measuring how teacher autonomy is associated with student learning and achievement. Neither the findings of the existing studies are always consistently evidencing the importance of teacher autonomy for student achievement. Also, due to the multifaceted nature of teacher autonomy, different studies explore autonomy conceptualized differently. Some studies explore this concept at the teacher-level, while others look at the school-level autonomy of school personnel, including teachers, in different domains and their implications for student achievement and equity.

The study by Gurganious (2017) examined the relationship between the teachers' perceptions in making decisions about teaching and student achievement in eighth-grade science at Florida schools under a high-stakes testing environment, i.e., the presence of a strong output-control mechanism. He found that this relationship was not statistically significant, i.e., there was no association between students' achievement and teachers' perceptions about their educational autonomy.

Like Gurganious (2017), Mills (2016) also explored the effects of teacher autonomy on teachers' perception of their instructional freedom. His smaller-scale study with the Texas school community found student performance in math and science negatively related to teacher autonomy. As the authors argue, this negative relationship

can be explained by the fact that various teacher and school-related factors were not included in the study. Understanding the relationship between teacher autonomy and student achievement would benefit from studies taking more holistic approaches. This means incorporating and sufficiently analyzing all factors that may affect this relationship, as Mills (2016) noted. It is also important to acknowledge that tests account only for a small part of students learning and education (Kennedy, 1999), considering the multidimensional nature of the purpose of education and socialization and subjectification components of the goals along with the qualifications-related purposes (see Biesta, 2015a; Biesta, 2015b).

Other studies demonstrated mixed findings about the relationship between teacher autonomy and student achievement. Tahirsylaj (2019) analyzed student achievement scores from PISA 2009 across 12 countries with so-called didactic and curriculum traditions. Firstly, his findings did not support his hypothesis about higher levels of teacher autonomy and teacher responsibility in Didactic countries compared to the so-called curriculum ones. He explained these disappearing differences between the two groups of countries by the increasing dominance of OECD-dictated policies and the quest of the countries to align their educational reforms with recipes by PISA and other international assessments. Secondly, within-country models of analysis provided evidence for the lack of predictive power of teacher autonomy and teacher responsibility in the majority of the participant countries.

Some existing empirical studies indicate that teacher and school-related autonomy is context specific and may have different implications for different groups of students or countries. In addition, there may be variations found in how different domains of

autonomy affect student achievement. For example, Cimentada (2020) studied the effects of school-level autonomy on student achievement using PISA data. He specifically looked at the variations in student achievement explained by academic (course content and textbooks) and administrative autonomy. The findings of his study indicated that academic autonomy impacted the achievement of different groups differently. Specifically, he found a positive association between autonomy over academic content and textbooks to be positively associated with the achievement of students in the top 10 % performers in mathematics (increased scores by 0.04 standard deviation). Contrary to this, he observed a negative association between autonomy in the same domain and achievement of the bottom 10 % performers (-0.06 standard deviation decrease). Cimentada (2020) also found that compared to academic autonomy, the explanatory power of administrative power was very weak for both groups of students.

Another study by Hanushek et al. (2013) examined the relationship between school autonomy and student achievement in the international PISA test during 2000-2009 in 42 countries. Similar to the study by Cimentada (2020), their work examined school autonomy in general. As school autonomy includes aspects of the scope of teacher decision-making along with other members of the school community, its findings are still relevant for this analysis. The central finding of Hanushek and his colleagues (2013) was that school autonomy improved student achievement in developed countries with a relatively high GDP and, by contrast, negatively affected student scores in less developed countries. These varying effects of autonomy were "most pronounced in decision-making on academic content, with some additional relevance for personnel autonomy and, less so, for budgetary autonomy" (Hanushek et al., 2013, p. 228). The authors attributed



undesirable consequences of autonomy in less developed countries to a lack of capacity and more potential for opportunistic behavior.

Based on the three-year case study of a primary school in England, Day (2020) provided additional case-specific evidence about the effects of teacher autonomy on student achievement. He found that these effects differed depending on how the teachers in the same school appropriated the reform initiatives. Teachers' use of individual autonomy created barriers rather than increased opportunities for improving student outcomes. The teachers with a strong moral purpose of educating their students acted collectively to change their practices for the best of their students. Other groups of teachers, lacking such moral purpose, "used the claim to individual professional autonomy as a way of excusing their resistance, despite the decline over three years in their students' progress and attainment" (Day, 2020, p. 258).

To conclude, the empirical evidence measuring the relationship between teacher autonomy and student achievement and the achievement gap is limited, as well as inconsistent. By examining the moderating effects of teacher autonomy as a country-level variable, this study will contribute to filling the existing gap in the scholarly literature. Specifically, analyses of the PIRLS 2016 data across 43 countries with various economic and democratic development and autonomy-related state regulations for teachers will shed light on the implications of input and output control mechanisms on teacher classroom performance and student achievement in reading.

## CHAPTER 3

### METHODOLOGY

In this section, I present the information about the methodology of the present study. I describe the secondary data I have used, my participants, and the analytical dataset. I also explain the process of creating all the measures for the study. I conclude the section by describing the statistical analysis tools I have employed to answer my research questions.

#### **Data**

To answer my research questions about the moderating role of teacher autonomy, I used secondary data from the PIRLS (Progress in International Reading Literacy Study) 2016, conducted by IEA (the International Association for the Evaluation of Educational Achievement). PIRLS 2016 is the fourth assessment conducted on a regular 5-year cycle. It assesses students' reading achievement after four years of schooling. The PIRLS 2016 examines student reading skills within two overarching purposes: (a) Reading for the literary experience; and (b) Reading to acquire and use information. Within each of these two purposes, it evaluated four types of comprehension processes: (a) Focus on and retrieve explicitly stated information; (b) Make straightforward inferences; (c) Interpret and integrate ideas and information; (d) Evaluate and critique content and textual elements (Mullis et al., 2015, pp. 18-22). PIRLS also collects comprehensive data on the educational contexts, organizational approaches, and instructional practices in participating countries (Mullis et al., 2017). I have selected PIRLS 2016 as the primary data source for my study as it allowed me to conduct a cross-sectional analysis with the participation of a high number of countries across the world. Also, the comprehensive

nature of its data permitted me to create the measures or use the readily available ones relevant for testing my hypotheses.

I used country-, student-, teacher-, parent- and school-level data from PIRLS 2016 Survey for this study. The country-level data drew from the curriculum questionnaire that was completed by the national coordinators of the participating countries. It gave information about the organization and content of the reading curriculum and other national policy issues, including types of input and output control concerning the reading curriculum. This data helped to operationalize autonomy in curriculum-making as freedom from interference or control in two directions: input control and output control (Martin et al., 2017).

This study also used two types of student-level data. As I mentioned above, the first is student achievement in reading, collected through reading tests administered to sampled students. The second type of data was taken from the student questionnaire given to each student after they completed their reading assessment. This questionnaire collected information about students' home environment, available books and resources, home language, attitudes towards reading, engagement in reading lessons, etc. (Martin et al., 2017).

Teacher-level data came from the teacher questionnaire, designed to collect information on teacher characteristics, such as education, years of experience, job satisfaction, etc. Most importantly, this questionnaire allowed unpacking classroom contexts for reading instruction, including instructional decisions and approaches used by the teachers (Martin et al., 2017). Given the focus of the study, teacher-level data was

central for measuring key-dependent and independent variables on teacher qualification and instructional practices.

A so-called home questionnaire completed by the guardians or parents of the participant students was critical for collecting students' family background information, such as their level of education or home conditions. Finally, school-level data came from a school questionnaire completed by school principals and department heads. They provided information about the demographic characteristics of their students, the availability of instructional resources, and the environment for learning in the schools (Martin et al., 2017).

### **Participants and Sampling**

PIRLS 2016 includes a nationally representative sample of the target population of students with four years of schooling, irrespective of their age. Students are randomly sampled in two stages. First, a sample of schools is drawn. In the second stage, one or more intact classes of students are selected from each of the sampled schools (LaRoche et al., 2017). Teachers do not make a representative sample of the entire population, as they were selected through the criteria of being teachers of randomly chosen students. PIRLS 2016 database includes student achievement data and student, home, teacher, school, and curricular background data for 50 countries and 11 benchmarking participants. Specifically, the PIRLS database makes available data from 346 852 students, 337 332 parents, 16 476 teachers, 12 124 school principals, and the national research coordinators of each country (Foy, 2018).

For the purpose of this study, I limited the data to 43 countries only (see Table 1). 11 benchmarking participants were not included in the study. In addition, I excluded

seven of the main participant countries: (a) Poland - For an unknown reason, information on Poland was missing from the curriculum questionnaire database; (b) United States – country representative for the United States did not provide answers to the curriculum questionnaire questions due to significant differences across the states; (c) South Africa, Kuwait, and Egypt – these three countries participated in the E-PIRLS version of the assessment only, (d) England – for unknown reasons, data for the student SES was missing from the original country data file; (e) Malta – Significant amount of information was also missing on the school-level data. As a result, I included 43 countries in the analysis. I should recognize that among 43 countries, several countries are represented twice. For example, Flemish and French Belgium have participated as separate units. Also, China is represented by two separate regions (Chinese Taipei and Macao). As a result, my final sample comprised 43 countries, represented by 245,696 students, 245 696 parents, 12,734 teachers, and 9,200 school principals (see Table 1). These 43 countries represented a range of geographical locations, socioeconomic, educational, and political systems, and policies about teacher autonomy and curriculum.

Table 1.

*PIRLS 2016 Countries Selected for the Study*

	Country Name	Number of Cases			
		School Principals	Students	Teachers	Parents / Caregivers
1	Australia	286	6,341	531	6,341
2	Austria	150	4,360	259	4,360
3	Azerbaijan	170	5,994	298	5,994
4	Bahrain	182	5,480	211	5,480
5	Belgium (Flemish)	148	5,198	277	5,198
6	Belgium (French)	158	4,623	254	4,623
7	Bulgaria	153	4,281	214	4,281

	Country Name	Number of Cases			
		School Principals	Students	Teachers	Parents / Caregivers
8	Canada	926	18,245	1,120	18,245
9	Chile	154	4,294	154	4,294
10	Chinese Taipei	150	4,326	176	4,326
11	Czech Republic	157	5,537	270	5,537
12	Denmark	185	3,508	186	3,508
13	Finland	151	4,896	296	4,896
14	France	163	4,767	284	4,767
15	Georgia	200	5,741	285	5,741
16	Germany	208	3,959	227	3,959
17	Hong Kong	138	3,349	150	3,349
18	Hungary	149	4,623	209	4,623
19	Iran	271	4,385	309	4,385
20	Ireland	148	4,607	219	4,607
21	Israel	159	4,041	159	4,041
22	Italy	149	3,940	217	3,940
23	Kazakhstan	172	4,925	234	4,925
24	Latvia	150	4,157	220	4,157
25	Lithuania	195	4,317	243	4,317
26	Macao	57	4,059	138	4,059
27	Morocco	360	5,489	386	5,489
28	The Netherlands	132	4,206	226	4,206
29	New Zealand	188	5,646	411	5,646
30	Northern Ireland	134	3,693	161	3,693
31	Norway	150	4,232	215	4,232
32	Oman	306	9,234	360	9,234
33	Portugal	218	4,642	318	4,642
34	Qatar	216	9,077	398	9,077
35	Russian Federation	206	4,577	213	4,577
36	Saudi Arabia	202	4,741	202	4,741
37	Singapore	177	6,488	354	6,488
28	Slovak Republic	220	5,451	334	5,451
39	Slovenia	160	4,499	253	4,499
40	Spain	629	14,595	678	14,595
41	Sweden	154	4,525	227	4,525
42	Trinidad and Tobago	151	4,177	195	4,177
43	United Arab Emirates	468	16,471	663	16,471
	<b>Total</b>	<b>9,200</b>	<b>245,696</b>	<b>12,734</b>	<b>245,696</b>

I accessed PIRLS 2016 public-use version data files in SPSS at IEA TIMSS and PIRLS international study center website:

<http://timssandpirls.bc.edu/pirls2016/international-database.html> in June 2021. The international database contains student achievement data files and student, teacher, school, and home background questionnaire data. I also accessed and downloaded the student-teacher linkage file. These files carried the following names at the IEA TIMSS and PIRLS international study center website:

- ACG\*\*\*R4: PIRLS school background data files
- ATG\*\*\*R4: PIRLS teacher background data files
- ASG\*\*\*R4: PIRLS student background data files
- ASH\*\*\*R4: PIRLS home background data files
- AST \*\*\*R4: PIRLS student-teacher linkage files

In addition to the files mentioned above, I accessed a school curriculum questionnaire data file available in Excel. I accessed and downloaded supplementary materials from the same section of the website: (a) Data almanacs with summary statistics for all PIRLS 2016 items and background variables, and (b) Codebook files describing all variables in the PIRLS 2016 international database.

After downloading student, teacher, school, and home background SPSS files for each country, I merged them into one dataset using the following steps: (1) I merged student background, teacher background, school background, and home background data files for each country into one combined file for each category; (2) I combined each country file for the teacher-student linkage file into one combined file; (3) I merged all

five combined files (student background, teacher background, school background, home background, and student-teacher linkage files) into one master dataset.

In addition to the SPSS files available at PIRLS international study center website, I created a new SPSS file that contained information on each country's input and output control and the type of autonomy. In this SPSS file, I included the following variables along with corresponding information for each country: country name, country ID, input control score, output control score, and type of autonomy. As a next step, I merged the SPSS file with control and autonomy-related country-level information to the master dataset.

PIRLS 2016 data files include 15 sampling and weighting variables, such as TOTWGT (Total student weight), SCHWGT (school weight), TCHWGT (teacher weight), etc. despite the availability of several weight variables, I only used the teacher weight variable, TCHWGT, for analysis, even with the student data (Foy, 2018). As explained in the PIRLS 2016 user guide for international databases, when using one all-encompassing merged data file, including student, home, teacher, and school background files, it is important to use TCHWGT as a weighting variable for all the analysis, including that of the student data. Foy (2018) explains this through the issues related to how TOTWGT (Total student weight variable) appropriates the student-teacher linkage file. Therefore, during the entire analysis process, I kept my master data file weighted by TCHWGT.

Due to missing information on some of the variables for some of the observations in the merged file, I deleted observations case-wise. After case wise deletion of the observations, the size of the sample was as follows: 173,666 students, 173,666 parents,



10,095 teachers, and 7,751 school principals. Application of TCHWGT weights to the data I have chosen for this study resulted in a final weighted analytic sample of 6,625,025 students, 6,625,025 parents, 460,726 teachers, 410,865 school principals.

### **Measures**

This section presents the measures I have used to answer my research questions. As I describe below, I have used some readily available variables. In addition, I have created several aggregate variables by combining two or more variables from student, teacher, home, and school data files (see Appendix 1). I calculated Cronbach's alpha to assess the internal consistency of these aggregated measures, i.e., the agreement among these variables in measuring the underlying dimension. In determining the acceptable alpha value, I used the level established by DeVellis (2021) and Kline (2005) - 0.7. However, I also considered a value close to 0.65 to constitute an acceptable level of internal consistency. To determine each item's contribution and to make a decision on the inclusion or exclusion of a variable from a construct, I relied on the values of Cronbach's alpha when that variable was excluded from calculations.

#### **Teacher Autonomy (as freedom from control)**

As I explained in the introduction chapter, for the purpose of this research, I conceptualize teacher autonomy as a decision-making space (Luthans, 1992, seen in Friedman, 1999), more specifically, the scope of action that state defines for teachers by its regulatory documents (Wermke, 2013) through granting them certain degrees of freedom from control. I differentiate between two types of control: input control, which targets resources for instruction, and output control, focused on performance measurement (Helgoy & Homme, 2007; Hopmann, 2003; Wermke & Forsberg, 2017).

Finally, as I explained below, I determine the type of country derived from the extent of teacher freedom from control (i.e., autonomy) based on the presence, absence, and combination of input and output control mechanisms.

PIRLS 2016 curriculum questionnaire includes items that allow me to measure autonomy as freedom from control from the national/regional government and identify two domains of control: input and output control. Specifically, the presence or absence of the input control mechanisms can be detected from the answers to questions about the prescriptiveness of the language/reading curriculum in the following areas: (a) goals and objectives, (b) instructional processes or methods, (c) materials, (d) assessment methods/activities. For measuring the output control, I used the items of the curriculum questionnaire focused on the evaluation of the implementation of the curriculum through the following mechanisms: (a) visits by inspectors, (b) research programs, (c) school self-evaluation, (d) national or regional examinations. After identifying items and related variables for each type of control, I assigned a score to each variable: Yes (i.e., presence of control) = 1, No (i.e., absence of control) – 0. Table 2 displays all the input and output control-related items and corresponding questions from the PIRLS 2016 curriculum questionnaire.

Table 2.

*Items for Operationalizing Teacher Autonomy as Freedom from Control*

<b>Item</b>	<b>Questions from the curriculum questionnaire</b>	<b>Response Scale</b>	<b>Scores assigned</b>
<b>1. Input control</b>			
1.1. National/regional control over goals and objectives	Does the language/reading curriculum prescribe goals and objectives?	Yes, No	Yes = 1 No = 0

<b>Item</b>	<b>Questions from the curriculum questionnaire</b>	<b>Response Scale</b>	<b>Scores assigned</b>
1.2.	National/regional over instructional processes or methods	Does the language/reading curriculum prescribe instructional processes or methods?	Yes, No Yes = 1 No = 0
1.3.	National/regional control over instructional materials	Does the language/reading curriculum prescribe materials (e.g., textbooks, instructional materials)	Yes, No Yes =1 No = 0
1.4.	National/regional over assessment methods and activities	Does the language/reading curriculum prescribe assessment methods/activities?	Yes, No Yes =1 No = 0
<b>2. Output control</b>			
2.1.	National/regional control through visits by inspectors	Is the language/reading curriculum implementation evaluated through visits by inspectors?	Yes, No Yes =1 No =0
2.2.	National/regional control through research programs	Is the language/reading curriculum implementation evaluated through research programs?	Yes, No Yes =1 No =0
2.3.	National/regional control through self-evaluation	Is the language/reading curriculum implementation evaluated through school self-evaluation?	Yes, No Yes =1 No =0
2.4.	National/regional control through national or regional examinations	Is the language/reading curriculum implementation evaluated through national or regional examinations?	Yes, No Yes =1 No = 0

Based on the explanations provided by the country representatives in the comments section of the questionnaire, I adjusted the score assigned to one of the input control items -national/regional control over goals and objectives - for Northern Ireland and Norway. In addition, I adjusted the score for the output control item - national/regional control through national or regional examinations – for Finland. Table 3 explains the rationale for each change.

Table 3.

*The Description for Input and Output Control Score Adjustment*

<b>Country</b>	<b>Description of the change to initially assigned scores</b>	<b>Comments provided by the country representatives in the curriculum questionnaire</b>
Northern Ireland	Although the country representative answered “Yes” to the question “What does the language/reading curriculum prescribe? Assessment methods/activities”, I changed the initially assigned score of 1 to 0 in response to the country representative's explanation in the comments section.	“The curriculum prescribes the learning objectives and outcomes and minimum content schools are required to cover – it prescribes a minimum only, schools have the flexibility to build upon this to best meet the needs of their pupils. The assessment arrangements prescribe the knowledge, understanding and skills pupils should be able to demonstrate at each key stage. The Council for Curriculum, Examinations and Assessment provides guidance materials for teachers. It is for schools to decide on pedagogies, materials and resources they wish to use to deliver the curriculum”.
Norway	Although the country representative answered “Yes” to the question “What does the language/reading curriculum prescribe? Assessment methods/activities”, I changed the initially assigned score of 1 to 0 in response to the explanation the country representative provided in the comments section.	“Schools and teachers have freedom when it comes to methods and materials. There is a mandatory assessment in Reading in Grade 2 and Grade 4”.
Finland	Although the country representative answered “Yes” to the question “How is the language/reading curriculum implementation evaluated? National or regional examinations”, I changed the initially assigned score of 1 to 0 given the explanation the country representative provided in the comments section.	“There is no extensive evaluation that encompasses all the schools but there are some national and international studies that help in evaluation.”

Another key concept for creating the measure of teacher autonomy, as freedom from control, is the degree of control in both directions: input and output. As Cribb and Gewirtz (2007) point out, the state may exercise blatantly coercive mechanisms of control, significantly diminishing the degree of freedom, or gentle coercive ones, allowing for greater freedom. For example, Hudson (2007) differentiates subtle and indirect forms of output control, such as self-evaluation, and rigid and direct forms of control, such as standardized testing organized by national-level bodies. Similarly, the subtleness or rigidity of input control depends on the areas the national curriculum prescribes diligently or leaves open for teachers' or schools' discretion (Cribb and Gewirtz, 2007). Based on these differentiations, I created the following criteria for identifying strong and weak input control forms: solo presence or absence of the prescription of the goals and objectives by the national curriculum. As a result, I defined input control as weak if the national curriculum prescribed only goals and objectives and strong if any other areas, such as materials, teaching, and assessment methods, are also prescribed in addition to the goals and objectives. For defining the degree of output control, I used the presence of the national and regional examinations as the main criterion. Therefore, I described output control as weak if the country used any or all the control mechanisms, such as visits by inspectors, self-evaluation, and research programs, but national and regional examinations. Correspondingly, I determined the type of output control as strong if the system used national or regional examinations (by itself or in combination with other mechanisms) to evaluate the curriculum implementation.

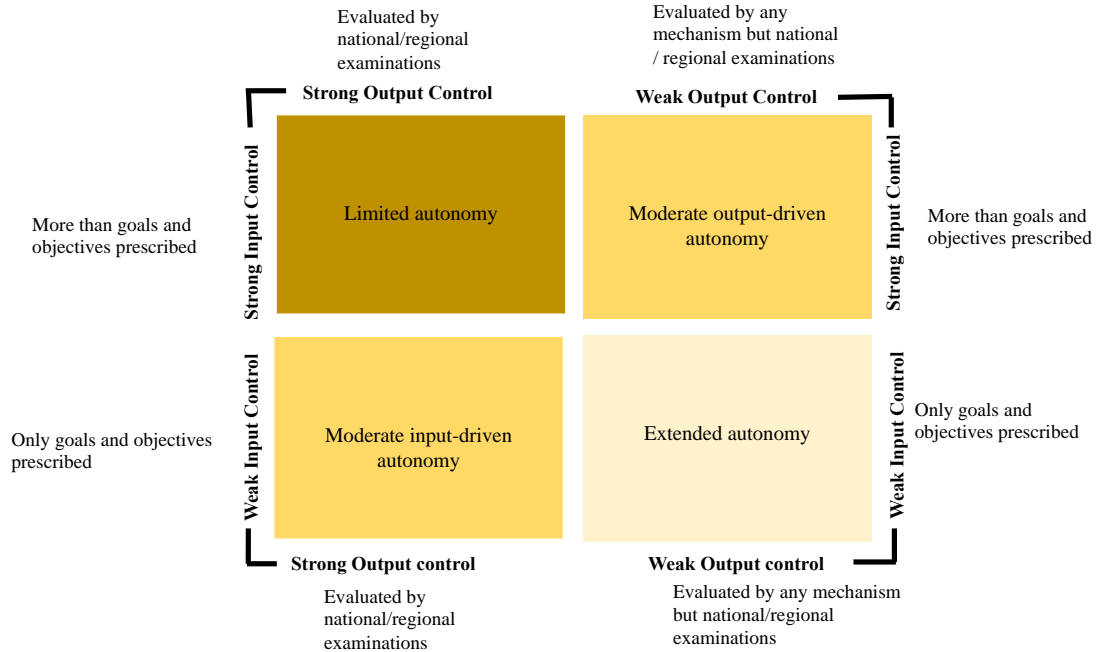
As a next step, I combined input and output control types (strong and weak) and created the following four categories of control: (1) Strong input and strong output

control; (2) Strong input and weak output control, (3) Weak input and strong output control, and (4) Weak input and weak output control. Finally, I determined the type of teacher autonomy freedom from control (i.e., autonomy) based on the presence and absence and combination of input and output control mechanisms:

- Presence of strong input and strong output control mechanisms - Limited autonomy;
- Presence of strong input and weak output control mechanisms – moderate output-driven autonomy;
- Presence of weak input and strong output control mechanisms –Moderate input-driven autonomy;
- Presence of weak input and weak output control mechanisms - Extended autonomy (See Figure 1).

Figure 1.

*Teacher Autonomy: Summarizing the Measure*



**Country’s Economic Development Level**

For measuring the country’s economic development level, I used the World Bank’s analytical classification of the world's economies based on estimates of Gross National Income (GNI) per capita for the 2016 calendar year (The World Bank, 2022). For the calendar year of 2016, the World Bank defined low-income economies those with a GNI per capita of \$1,005 or less; lower-middle-income economies are those with a GNI per capita between \$1,006 and \$3,955; upper-middle-income economies are those with a GNI per capita between \$3,956 and \$12,235; high-income economies are those with a GNI per capita of \$12,236 or more (World Bank Data Team, 2016).

## **Country's Democratic Development Level**

The Economist Intelligence Unit's (2017) Democracy Index provides a snapshot of the state of democracy worldwide. The Democracy Index is based on five categories: electoral process and pluralism; civil liberties; the functioning of government; political participation; and political culture. Based on their scores on a range of indicators within these categories, each country is classified as one of four types of regimes: full democracy, flawed democracy, hybrid regime, and authoritarian regime. According to the Economist Intelligence Unit (2017) Report, full democracies have a political culture that leads to the flourishing of democracy, independent media, and judiciary system. In flawed democracies, fundamental civil liberties are respected; however, the overall political culture is relatively underdeveloped. Hybrid regimes tend to have problems of unfair elections, corruption, pressure on opposition and media, and low political participation. Finally, authoritarian regimes are either dictatorships or have only formal institutions of democracy. Media is not free, and civil liberties are abused regularly.

## **Achievement Gap in Reading**

Approaches to measuring the achievement gap vary across educational studies. Some scholars study the achievement gap by students' socioeconomic background (Chmielewski, 2019). Others use a more focused approach and base their approach on parental education or income. For example, Reardon (2011) measured the achievement gap as the difference in test scores between students from families at the 90th and 10th percentiles of the income distribution. Instead, Saw (2016) prioritized parental education as the basis for estimating the achievement gap. Due to the modest correlation between parental income and education in the Norwegian context, Sandsør et al. (2023) estimated



achievement gaps by education and income. I based the achievement gap on student performance, i.e., the difference between the average scores of students in the 75<sup>th</sup> and 25<sup>th</sup> percentiles of PIRLS score distribution. In addition, to provide equity-related conclusions, I analyzed how factors representing teacher autonomy affect the relationship between socioeconomic background and student achievement (research question 6).

### **Share of Economically Disadvantaged Students at School**

In addition to the student's socioeconomic background, school composition by student socioeconomic background is also associated with student achievement (Rumberger & Palardy, 2005; Willms, 2010). In literature, scholars often operationalize school composition by SES as the aggregate of the social-economic background of the students attending the school (Perry & McConney, 2010). PIRLS data allows measuring school composition by SES as the share of students with low socioeconomic backgrounds. School principals or department heads responded to the survey question about the approximate percentage of students in their school who come from economically disadvantaged homes on a four-level scale: (1) 0-10 %; (2) 11-25 %, (3) 26-50 %, (4) More than 50 %.

### **Instruction Affected by Reading Resource Shortages**

The availability of quality school and classroom resources for instruction is key to ensuring quality instruction for students. Results from large-scale assessments show a strong correlation between principals' perceptions of the shortages of school resources and student achievement (Lay & Chandrasegaran, 2016; Mullis et al., 2015). If there is a lack of such a correlation, it may be due to the improper use of these resources for teaching and learning rather than the insignificance of the availability of such resources

(Miningou et al., 2022). The availability of resources as a pre-condition for closing an achievement gap is particularly crucial for non-OECD countries (Yang & Lee, 2022). PIRLS 2016 has a readily available scale called Classroom Instruction Affected by Reading Resource Shortages (RRS). The scale measures the availability of both general school resources and resources for reading instruction. As described in Appendix 1, the scale uses items for two types of resources: (a) general school resources and resources for reading instruction. PIRLS 2016 has created two versions of this measure. First, after transforming raw scores into scale scores, it created a continuous variable with a minimum score of 3.55, maximum score of 14.49. Based on the scale scores, PIRLS also created a categorical measure as well with the following scale: (1) Not affected, (2) Somewhat affected; (3) Affected a lot (details about the construct can be seen in Martin et al., 2017).

### **Teacher Qualification**

In teacher-related literature, teacher qualification, along with other teacher characteristics, is seen as one of the key categories of input (i.e., the resources teachers bring to the classroom, Goe, 2007) to teacher quality. Qualification is usually represented through multiple categories. For example, in her comprehensive framework for teacher quality, Goe (2007) identifies the following forms of qualification: teachers' coursework, grades, subject matter education, degrees, test scores, experience, certification, and credentials, as well as evidence of participation in continued learning such as internships, induction, supplemental training, and professional development. Quite similarly, Zuzovsky (2009) identifies the following key components of teacher qualification: teacher's formal education, teacher education in the subject matter of teaching, teacher

education in pedagogical studies, duration of pre-service education, certification and licensing status, years of experience, participation in professional development activities.

For this study, I selected four major categories to operationalize teacher qualification: formal education, teaching experience, formal professional development, and school-based professional development. In education research, the level of teacher education and years of experience is one of the most widely used proxies of teacher qualification (Abe & Adu, 2014; Hanushek et al., 2005; Ladd, 2008; Podolsky et al., 2019; Wiswall, 2013). I selected participation in two forms of professional development (formal and school-based) as additional categories of teacher qualification due to the evidence of their power to improve teacher effectiveness and positively influence student achievement (Blank & Alas, 2010; Egodawatte, 2011; Harris and Saas, 2007; Johnson, 2003; Meirink, 2007; Leana & Pil, 2006; Slavit et al., 2011).

### ***Teacher Education***

I measured teacher education as the highest level of education attained by the teacher. I used the following Likert-scale items from the teacher background questionnaire - What is the highest level of formal education you have completed (Item G4)? The scale for this item was composed of seven levels: (1) Did not complete upper secondary education—ISCED Level 3; (2) Upper secondary education—ISCED Level 3, (3) Post-secondary, non-tertiary education—ISCED Level 4; (4) Short-cycle tertiary education—ISCED Level 5; (5) Bachelor's or equivalent level—ISCED Level 6; (6) Master's or equivalent level—ISCED Level 7; (7) Doctor or equivalent level—ISCED Level 8. I recoded an original teacher education variable from a 7-level scale to a 2-level scale with the following distribution: 1- Bachelor's or lower (including the following:

ISCED Level 3, ISCED Level 4, ISCED Level 5, ISCED Level 6) 2 – Master’s or above (including the following: ISCED Level 7, ISCED Level 8).

### ***Teaching Experience***

I defined teaching experience as the number of teaching years reported by the teacher. I used the following open-ended item from the teacher questionnaire - By the end of this school year, how many years will you have been teaching altogether? I have included two versions of the teaching experience measure. First, I kept the original variable in a continuous form (minimum – 0 years, maximum - 60 years). Second, I also created a categorical version of the variable on teaching years. I reverse-coded the original variable using the following 4-level scale: (1) less than 5 years, (2) at least 5 years but less than 10 years, (3) at least 10 years but less than 20 years, and (4) – 20 years or more.

### ***Formal Professional Development***

I operationalized teachers’ formal professional development as the intensity of participation in various formalized professional development opportunities. PIRLS 2016 teacher questionnaire included one item that captures the intensity of teacher participation in formal professional development: “In the past two years, how many hours in total have you spent in formal professional development (e.g., workshops, seminars, lesson studies, etc.) that dealt directly with reading or teaching reading (e.g., reading theory, instructional methods)?”. I kept the original 5-level scale the item and corresponding variable had in the questionnaire and the teacher database: (1) None, (2) Less than 6 hours, (3) 6-15 hours, (4) 16-35 hours, (5) More than 35 hours.

### ***School-based Professional Development***

I measured teachers' school-based professional development by the intensity of their participation in various school-based professional development activities. The creation of the school-based teacher professional development measure was informed and guided by Little's (1990) conceptualization of teacher collaboration at the school level, including storytelling and scanning for ideas, aid and assistance, sharing, and joint work. The following four items from the teacher's questionnaire allowed for measuring the above-mentioned forms of school-based teacher professional development through peer collaboration: (a) Sharing what the teacher has learned about teaching experiences; (b) Observing another classroom to learn more about teaching; (c) Working together to improve how to teach a particular topic; (e) Working with teachers from other grades to ensure continuity in learning. All four items were preceded by the following question: How often do you have the following types of interactions with other teachers? The question employed a 4-level scale of measurement with (4) Never or almost never, (3) Sometimes, (2) Often, and (1) Very often. The measure was reliable as the selected items yielded an alpha of 0.8. I created two versions of the measure: continuous and categorical. For creating a continuous variable, I reverse-coded the measuring scale of all items as follows: 1 - Never or almost never, 2 - Sometimes, 3 - Often, and 4 - Very often. After reverse coding, I computed a continuous aggregated variable by adding all the values of each item contributing to the aggregate variable (minimum – 4 points, maximum – 16 points). I also created a categorical version of the school-based teacher professional development variable. I recoded the continuous variable I had computed into a 4-level scale using the following thresholds: (1) Never or almost never – 4 to 6 points,

(2) Sometimes – 7 to 10 points, (3) Often – 11 to 14 points, (4) Very often – 15 to 16 points.

### **Classroom practices**

Classroom practices or instructional activities are key to student achievement, as most teaching and learning processes occur in school classrooms (Mullis et al., 2017). According to Goe's (2007) teacher quality framework, teacher practices, i.e., planning, instructional delivery, classroom management, and interactions with students, makes a critical process-related component of the teacher quality framework, and together with teacher qualifications and teacher characteristics (defined as inputs), contributes to teacher effectiveness (i.e., outcome), that is determined by growth in student learning. Although students perform learning activities, the teacher's role is significant as they are facilitators or regulators of this process. Therefore, I defined teacher classroom practices as teachers' actions "to initiate, monitor, focus, modify, control, influence, direct, correct, maintain) students' behaviors while completing their learning activities" (Den Brok et al., 2004, p. 426).

As described below, I have selected four sets of teacher classroom practices. First, given the focus of the study on student achievement in reading, I selected teacher use of reading comprehension strategies as one of the four sets of classroom practice. Two other practices are more generic and key for differentiated instruction, that "allows all students to access the same classroom curriculum by providing entry points, learning tasks, and outcomes tailored to students' learning needs" (Watts-Taffe et al., 2012, p. 304). These practices are teacher use of individualized instruction and content differentiation. Finally, given the importance of student autonomy for student learning and achievement (Orakci

& Gelisli, 2019; Stefanou et al., 2004;), I also included teacher use of autonomy-supportive strategies as the fourth set of key classroom practices.

### ***Teacher Use of Reading Comprehension Strategies***

For this study, I operationalize teacher use of reading comprehension strategies as the frequency of teacher-reported activities aimed at developing student skills of predicting, making connections, summarizing, inferring, understanding text structure, discussing, and analyzing (Duke et al., 2021; Küçükoğlu, 2013). Question 12 of the teacher questionnaire (How often do you ask the students to do the following things to help develop reading comprehension skills or strategies?) includes nine items that capture the frequency of the use of effective strategies for helping students to build reading comprehension skills: having students (a) Locate information within the text; (b) Identify the main ideas of what they have read; (c) Explain or support their understanding of what they have read; (d) Compare what they have read with experiences they have had; (e) Compare what they have read with other things they have read; (f) Make predictions about what will happen next in the text they are reading; (g) Make generalizations and draw inferences based on what they have read; (h) Describe the style or structure of the text they have read; (i) Determine the author's perspective or intention. All the selected items had the same Likert scale: (1) every day or almost every day, (2) once or twice a week, (3) once or twice a month, (4) never or almost never. The scale had a high-level internal consistency, as determined by a Cronbach's alpha of 0.898.

Using the items mentioned above, I created a continuous variable indicating the frequency of the teacher's use of reading comprehension strategies. I first reverse-coded the PIRLS 2016 original variables for each item, i.e., assigned the highest score of 4 to

the most frequent use of the strategy – every day or almost every day and changed all other score points accordingly. Then I added all nine scores and received a continuous variable with a range of points from 9 to 36.

### ***Teacher Use of Reading Content Differentiation***

I measured teacher use of reading content differentiation as the frequency of teacher-reported practices of offering varied text and resource materials to students (Heacox, 2018; Tieso, 2005) based on students' readiness level and interest (Tomlinson, 2005). The following two items from the teacher's questionnaire captured the proposed definition of reading content differentiation: (1) Provide reading materials that match the students' interests; (b) Provide materials that are appropriate for the reading levels of individual students. Both items were preceded by the following question (#11): How often do you do the following in teaching reading to this class? Also, both items use the following scale: Every or almost every lesson; About half the lessons; Some lessons; Never. According to the results, the construct had acceptable reliability as the selected variables yielded an alpha of 0.656. Using these items, I created a continuous version of the measure. I first reverse-coded both original variables into the following: 1 - Never or almost never; 2 – Sometimes; 3 – Often; 4 - Always or almost always. Then I created their sum, resulting in scores ranging from 2 to 8, with the higher number indicating more frequent use of reading content differentiation.

### ***Teacher Use of Autonomy-Supportive Strategies***

To conceptualize autonomy-supportive strategies, I consulted Alley's (2019) analysis of autonomy-supportive teacher behaviors that include: (1) giving students choices, (2) providing informative feedback, (3) letting them decide how they want to



learn; (4) encouraging the expression of criticism and independent opinions; (5) encouraging self-initiation, (6) minimizing the use of controls, and (7) encouraging collaborative learning. I selected seven items from the teacher's questionnaire to capture the teacher-reported use of autonomy-supportive strategies described above. The first set of items was preceded by the question: How often do you do the following in teaching reading to this class? I used the following five items: 1) give students time to read books of their choosing; 2) give individualized feedback to each student; 3) encourage students to challenge the opinion expressed in the text; 4) encourage students to develop their understandings of the text; 5) encourage student discussions of texts. The sixth item was part of the question - When you have reading instruction and/or do reading activities, how often do you organize students in the following ways?: Students work independently on an assigned plan or goal. The final item of the measure is the following: Talk with each other about what they have read. All items except the one on working independently on an assigned plan or goal were measured on the following scale: every or almost every lesson, about half the lessons, some lessons, never. That remaining item used the following scale: Always or almost always, Often, Sometimes, Never. The aggregate variable was reliable, with an alpha of 0.702.

I created a continuous version of the measure using the seven above-mentioned items. First, I reverse-coded all the variables so that the higher number assigned would express the higher frequency of the strategy use. Then, I computed the sum of the values of all seven variables, resulting in scores with a range from 7 to 28 (with the higher number indicating more frequent use of autonomy-supportive strategies).

### ***Teacher Use of Individualized Instruction***

For this study, I define individualized instruction as “providing individualized attention (in the form of specific instruction or feedback) to specific students” within their regular classroom instruction (Tetzlaff et al., 2022, p. 2). Accordingly, I operationalize teacher use of individualized instruction as the frequency of teacher-reported individualized attention in the form of specific instruction (How often do you do the following in teaching reading to this class? I use individualized instruction for reading). The item used a 4-level Likert scale: Every or almost every lesson; About half the lessons; Some lessons; Never). First, I reverse-coded the variable and then recoded it into a two-level scale: 1 Never or almost never, or sometimes, 2 - Often, always, or almost always.

### **Classroom Size**

Classroom size, i.e., the number of students in the classroom, has implications for teaching, student attentiveness, and social relations. For example, classroom size may affect teacher task time with students, teacher/individual support for learning, classroom management and control, and the depth of curriculum coverage (Blatchford et al., 2003). Therefore, I have included classroom size as a control variable in this study. Specifically, I have used a question from the student questionnaire: “How many students are in this class?”. In all my analyses, I have used it as a continuous variable (minimum – 1, maximum – 89).

### **Instruction Limited by Student Attributes**

PIRLS 2016 has a readily available scale called Instruction Limited by student attributes (SLI) to measure teachers’ perception of the limitations caused by various

student attributes. The scale uses items related to students suffering from prerequisite knowledge or skills, lack of basic nutrition, not enough sleep, as well as absent, disruptive, or unimpressed students. Finally, the scale also accounts for students with mental, emotional, or psychological impairment. PIRLS 2016 has created two versions of this measure. First, after transforming raw scores into scale scores, it created a continuous variable with a minimum score of 3.09 and maximum score of 14.37. Based on the scale scores, PIRLS also created a categorical measure as well with the following scale: (1) A little, (2) Some, (3) A lot (details about the construct can be found in Martin et al., 2017).

### **Student Achievement in Reading**

I measured student achievement in reading by scores reflecting their performance in the following four processes for two purposes of reading - reading for literary experience and reading to acquire and use information: (1) focus on and retrieve explicitly stated information; (2) make straightforward inferences; (3) interpret and integrate ideas and information; (4) evaluate and critique content and textual elements. Specifically, I computed the average of five overall plausible value items (readily available in the PIRLS 2016 student database). The average student scores range from 83.2 to 785.3.

### **Student Socioeconomic Background / Home Resources for Learning (HRL)**

A student's socioeconomic background is one of the most frequently used indicators to explain differences in student achievement. Socioeconomic background typically relates to an individual's (or family's) status within a given social hierarchy (Hattie, 2009, seen in Rutkowski & Rutkowski, 2013). Parents' income, occupation, and education are the most widely used parameters for measuring students' SES (Sandsør et

al., 2023). PIRLS 2016 has a readily available scale called Home Resources for Learning (HRL) to measure students' socio-economic backgrounds. It uses the following five items from student and parent questionnaires: (a) the number of books at home, (b) the number of home study support, (c) the number of children's books in the home, (d) the highest level of education of either parent, (e) highest level of occupation of either parent. PIRLS 2016 has created two versions of this measure. First, after transforming raw scores into scale scores, it created a continuous variable with a minimum score of 4.03, and maximum score of 14.8. Based on the scale scores, it created a categorical measure as well with the following scale: (1) Few resources, (2) Some resources; (3) Many resources (details about the construct can be seen in Martin et al., 2017).

### **Student Gender**

Analysis of the relationship between gender and student achievement in reading shows that the former plays a role, and girls quite significantly outperform boys (Voyer & Voyer, 2014). Therefore, I have included gender as a control variable. Specifically, I have used a question from the student questionnaire "Are you a girl or a boy?" (Response scale: Girl, Boy). I recoded the original variables as follows: 0 – Boy, 1 – Girl.

### **Data Analysis**

For analyzing PIRLS 2016 data and conducting all statistical tests, I used the statistical software platform SPSS (version 27.0) for Mac. Also, for analyzing curriculum questionnaire data that was initially available in Excel format, I used Microsoft Excel for Mac (version 16.60).

### **Tests for Comparing and Testing Means**

Given the descriptive nature of research questions 2 and 3, I used several statistical tests to compare the distribution of the main variables representing different dimensions of teacher qualification and teacher practices and the differences in the average values of these variables by the type of country and student characteristics.

***One-way analysis of variance (One-way ANOVA)***

I used a one-way analysis of variance (ANOVA) to determine whether there were any statistically significant differences between the means of two or more independent groups. Specifically, I performed this test to understand whether student achievement in reading differed: (1) across the groups of countries with different forms of teacher autonomy in curriculum implementation (Research question two), (2) by the country's economic and democratic development level in each of the four groups by type of teacher autonomy, (3) by the student socio-economic background (same as home resources) in each of the four groups by type of teacher autonomy (Research question two), (4) by share of economically disadvantaged students at school in each of the four groups by type of teacher autonomy (research question 2). In addition, I used one-way ANOVA to examine differences in the frequency of the classroom use of reading comprehension, content differentiation, and autonomy-supportive strategies across the systems with different types of teacher autonomy (research question 3). Using the same test, I also examined how teachers' application of selected classroom practices varied by the share of economically disadvantaged students at school teachers' perceptions of student-related attributes limiting their instruction (research question 3). Due to the categorical nature of the individualized instruction variable, I used a chi-square test to examine the similar relationship between teacher autonomy and the use of individualized instruction.

For all four analyses, my null hypothesis was that there was no significant difference between the group means ( $\mu$  = group mean and  $k$  = number of groups):  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ . For interpreting the results of one-way ANOVA, I checked the values for mean, standard deviation, and confidence intervals for the dependent variable across all groups of the independent variable. For rejecting or accepting the null hypothesis, I examined the significance value (alpha set at 0.05) and F Statistics. In addition, as one-way ANOVA is an omnibus test statistic and only provides information on the statistical significance as a whole (i.e., between at least two groups), I also conducted a follow-up post hoc test using the Bonferroni adjustments. Bonferroni's multiple independent samples t-test helped me determine which groups for each factor had statistically significant differences in the means of the dependent variable groups. Finally, I calculated eta squared ( $\eta^2 = SS_{\text{effect}} / SS_{\text{total}}$ ) to measure effect size and checked the Point Estimate value. I used the following thresholds for interpreting the magnitude of the effect size based on the Point Estimate value: small - 0.01, medium - 0.06, and large - 0.14 (Miles & Shevlin, 2001).

### ***Analysis of Variance (Two-way ANOVA)***

In addition, I conducted the two-way ANOVA to examine whether a two-way interaction effect existed between two independent variables on a dependent variable. Specifically, for research question 2, I estimated the joint effect of teacher autonomy and (1) the country's economic development level, (2) the country's democratic development level, (3) student socioeconomic background, (4) the share of economically disadvantaged students at school on student achievement in reading. Similarly, for research question 3, I tested the joint effect of teacher autonomy and (1) the share of

economically disadvantaged students at school, (2) the teacher's perception of classroom instruction limited by student attributes. My null hypothesis for all analyses was that my two independent variables had no significant interaction effect on the dependent variable. I rejected the null hypothesis based on the significance of the interaction coefficient.

### ***Independent Samples t-Test***

I used an Independent Sample t-Test to analyze the distribution of average student scores in reading by students in the top and bottom quartiles (research question 2), as well as by the level of economic and democratic development (when only two groups were available to compare). The null hypothesis for the independent t-test was that the population means from the two unrelated groups are equal:  $H_0: \mu_1 = \mu_2$  I assessed the value of significance (alpha set at 0.05) to accept or reject my null hypothesis.

### ***Moderation Analysis –Multiple Regression with Interactions***

I conducted a moderation analysis to examine how teacher autonomy moderated the relationship (1) between teacher qualification characteristics and classroom practices (research question 4), classroom practices and student achievement (research question 5), and student socio-economic background, classroom practices, and student achievement (research question 6). As part of research question 6, it also allowed me to test the hypothesis about the moderating role of teacher classroom practices. This analytical approach was appropriate for my study as moderation analysis is designed to “measure and test the differential effect of the independent variable on the dependent variable as a function of the moderator” (Baron & Kenny, 1986, p. 1174). There are other reasons (highlighted by Memon et al., 2019) that justified the appropriateness of selecting the moderation analysis for my studies. First, it allowed me to test some of the previous

inconclusive findings in the literature about the implications of teacher autonomy for educational systems. Also, I identified and chose teacher autonomy, as a primary moderator variable of interest, based on the considerable scholarly literature and theoretical support.

For assessing a moderation effect of teacher autonomy (research questions 4, 5, and 6), as well as selected four classroom practices (research question 6), I used a regression model with two-way and three-way interaction allowing to explain whether these variables were capable of affecting the strength and/or direction of the relationship between predictor and dependent variables (Andersson et al., 2020). As described by Dawson (2014), the regression equation for moderation with two-way interactions is as follows:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon [2]$$

This equation includes a dependent variable,  $Y$ , a predictor variable,  $X$ , a moderator variable,  $Z$ , and the interaction term  $XZ$  created by multiplying  $X$  and  $Z$ . Also, in the equation where  $\beta_0$  is the intercept,  $\beta_1$  is the coefficient of, and  $\varepsilon$  is the residual (error term). The inclusion of the second moderator,  $W$ , extends the equation as follows:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 W + \beta_4 XZ + \beta_5 XW + \beta_6 WZ + \beta_7 XZW + \varepsilon [5]$$

As shown, the extended equation for the three-way interactions includes the main effects of the predictor variable and the two moderators and the three two-way interaction terms between each pair of variables. Finally, it also includes the three-way interaction term (Dawson, 2014). For my analysis, I used variables for  $X$  and  $Y$  in their raw form.

I used five moderator variables for my study: teacher autonomy and four classroom-practice-related variables: reading comprehension strategies, reading content



differentiation, autonomy-supportive strategies, and individualized instruction. Teacher autonomy, a categorical variable, served as a primary moderator for research questions 4 and 5 and a secondary moderator – for research question 6. As testing for moderation with categorical variables with more than two categories requires special treatment (Dawson, 2014), I created dummy variables for three categories: limited, moderate output-driven, and moderate input-driven and entered them into the regression analysis. I used the fourth category, extended autonomy, as a reference category. I entered the three dummy variables into the regression as the first and second moderator variables. In addition, I created two-way (for research questions 4, 5, and 6) and three-way interaction terms (for research question 6) for each of those three dummy variables.

I used a hierarchical approach to test the hypothesis of all three questions pertaining to the moderation analysis, i.e., I entered the variables in different steps. The primary rationale for this decision was to allow for the computation of the incremental change in  $R^2$  that could be attributed to the inclusion of additional variables (Dawson, 2014). Also, it allowed me first to hold constant for all control variables. Table 1 summarizes the sequence of the steps of entering variables for all relevant research questions:

Table 4.

*Steps of Entering Variables in Linear Regression*

<b>Research Question</b>	<b>Model</b>	<b>Variables</b>
Research question 4:	Model 1	All control variables
	Model 2	Predictor / independent variable
	Model 3	Moderator variable
	Model 4	Two-way Interactions term
	Model 1	Student-level control variables

Research Question	Model	Variables
Research question 5:	Model 2	School-level control variables
	Model 3	Country-level control variables
	Model 4	Teacher-level control variables
	Model 5	Predictor / independent variable
	Model 6	Moderator variable
	Model 7	Set of two-way Interactions terms (predictor * dummy moderator)
	Research question 6:	Model 1
Model 2		Predictor / independent variable
Model 3		First moderator variable
Model 4		Two-way Interactions term (predictor * moderator)
Model 5		Secondary moderator variable
Model 6		Set 1 of two-way interaction terms (predictor * dummy secondary moderator) Set 2 of two-way interaction terms (first moderator * dummy secondary moderator)
Model 7		Set of three-way interaction terms (predictor * first moderator * secondary moderator)

I used standardized coefficients to interpret the findings. For moderation with two-way interactions, I first examined the significance of the interaction term. If it was statistically significant, I concluded that  $Z$  was a statistically significant moderator of the linear relationship between  $X$  and  $Y$ . For dummy variables as moderators, the statistically significant coefficient on interaction indicated that the moderating variable changed the magnitude of the difference in the outcome between the groups represented by independent variables. I also examined the coefficients  $\beta_1$  and  $\beta_2$  to determine whether there was any main effect of predictor ( $X$ ) and moderator ( $Z$ ) variables, respectively, and the value of  $\beta_3$  for determining the presence or absence of moderation. The positive sign of the interaction term showed that the relationship got stronger (in the case of categorical moderators – in relation to the reference dummy variable). I used similar judgment for interpreting the findings of three-way interactions. I examined if the coefficient  $b_7$ , i.e., three-way interaction term, was significant, i.e., had a statistically significant moderating

effect on another moderator ( $Z$ ) influencing the relationship between  $X - Y$  relationship.

In addition, as mentioned above, I also assessed the size of the R-squared change for each model involving the main effect and interaction variables.

## CHAPTER 4

### RESULTS

#### **Descriptive Statistics**

In this descriptive overview, I have included all variables of the study. As Table 5 depicts, about half of the 43 countries were categorized as countries with limited teacher autonomy (44 %). The second largest group had moderate input-driven teacher autonomy (26 %). Finally, countries with moderate output-driven and extended teacher autonomy made the smallest share (14 % and 16 %, correspondingly). Also, more than 80 % of the 43 countries were high-income economies. The other two groups had significantly lower share (lower-middle-income - 5 % and upper-middle-income – 14 %). Based on the democratic development level, approximately 70 % of the countries were either flawed or full democracies. Another largest group was the one with authoritarian regimes (26 %). Finally, hybrid regimes were the smallest group (5 %).

There were variations in the distribution of schools by the share of economically disadvantaged students. The largest share of schools in these 43 countries had only up to 10 % of students with low socioeconomic backgrounds. Approximately the same (27 %) was the share of those schools where 50 % or more students came from disadvantaged backgrounds. Schools were differently affected by reading resource shortages. Approximately 70 % were somewhat affected, while only 7 % experienced a high shortage. When measured as a continuous variable, the mean resource shortage made 9.6 with an *SD* of 1.8.

In addition to country-level and school-level variables, Table 5 includes an overview of the teacher-level measures. As shown, half of the teachers in all 43 countries

had a medium-level education, i.e., a BA degree or equivalent. Those with lower (less than BA) and higher (MA or higher) were represented with the same share (approximately 25 %). The average teaching experience was 18.3 years, with an *SD* of 11.0. Statistics related to formal professional development revealed that more than 20 % of teachers never participated in any related activities. The same share of teachers reported participation equaling less than 6 hours. The rest of the teachers participated for 6 hours or more. Interestingly, the percentage of teachers who had never participated in professional development was significantly lower for the school-based aspect of professional learning (4.6%). Also, more than 85 % of the teachers sometimes or often engaged in such activities. When measured by a continuous variable, the mean score of the intensity of participation was 10.9, with an *SD* of 2.7.

Statistics related to the frequency of teacher classroom practices show that the mean score for reading comprehension strategies was 29.2, with an *SD* of 5.2. The same values for the other two practices were as follows: reading content differentiation –  $M = 5.7$ ,  $SD = 1.6$ , and autonomy-supportive strategies -  $M = 21.6$ ,  $SD = 3.5$ . In addition, almost 70 % of teachers reported that they never or sometimes used individualized instruction with their students. Finally, the mean score of classroom size was 22.8, with an *SD* of 7.7.

Table 5 also includes important demographic and academic achievement information about students. Girls were half of the sample (49.5%). The mean score for students' socioeconomic status (measured as availability of home resources) was 9.8, with an *SD* of 2.0. A descriptive analysis of the categorical version of the socioeconomic variable shows that most students had some resources (73 %). Only 15.5 % of them had

access to many resources. Finally, the mean score in reading for students in all 43 countries was 513.2, with an *SD* of 100.3.

Table 5.

*Descriptive Statistics*

<b>Variable</b>	<b>Mean (SD)</b>	<b>Range</b>	<b>%</b>
<i>Country-level</i>			
Teacher autonomy			
Limited			44.0
Moderate output-driven			14.0
Moderate input-driven			26.0
Extended			16.0
Country's economic development level			
Lower-middle-income			5.0
Upper-middle-income			14.0
High-income			81.0
Country's democratic development level			
Authoritarian regime			26.0
Hybrid regime			5.0
Flawed democracy			39.0
Full democracy			30.0
<i>School-level</i>			
Share of economically disadvantaged students at school			
0-10 %			30.1
11 – 25 %			24.5
26-50 %			18.3
More than 50 %			27.1
Instruction affected by reading resource shortages (index)			22.7
Not affected			69.7
Somewhat affected			7.7
Affected a lot			
Instruction affected by reading resource shortages (Scale)	9.6 (1.8)	3.5-14.5	
<i>Teacher-level</i>			
Teacher education			
Less than BA			24.5
BA or equivalent			50.9
MA or higher			24.6
Teaching experience	18.3 (11.0)	0-60	
Formal professional development			
None			21.4
Less than 6 hours			20.9
6-15 hours			20.5
16-35 hours			15.8

<b>Variable</b>	<b>Mean (SD)</b>	<b>Range</b>	<b>%</b>
More than 35 hours			21.4
School-based professional development (Index)			
Never or almost never			4.6
Sometimes			41.1
Often			44.6
Very often			9.7
School-based professional development (Scale)	10.9 (2.7)	4-16	
Use of reading comprehension strategies	29.2 (5.2)	9-36	
Use of reading content differentiation	5.7 (1.6)	2-8	
Use of autonomy-supportive strategies	21.6 (3.5)	7-28	
Use of individualized instruction (Never/sometimes)			57.9
Classroom size	22.8 (7.7)	1-89	
<b><i>Student-level</i></b>			
Student achievement in reading (score)	513.2 (100.3)	83.2 - 785.3	
Student socio-economic background (Home resources for learning)			
Few resources			11.5
Some resources			73.0
Many resources			15.5
Student socio-economic background (Home resources for learning)	9.8 (2.0)	4.03-14.8	
Student gender (girl)			49.5

## **Main Findings**

### **Overview of the Countries by Teacher Autonomy in Curriculum Implementation**

In this section, I present the findings of the country-level data analysis of PIRLS 2016 countries by their input and output control mechanisms and, eventually, by type of teacher autonomy as freedom from the national level control from the national/regional government in curriculum implementation. Table 1 describes a summary of the responses by the country representatives to the eight items I selected for measuring teacher autonomy as the freedom from national control (including both input and output control components), as well as the type of input and output control assigned to each country.

As shown in Table 6, 18 out of 43 participant countries (i.e., 42 %) exercise weak input control over their teachers, i.e., their national curriculum prescribes only goals and

objectives and leaves the rest at the discretion of teachers: Australia, Belgium (Flemish), Belgium (French), Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Lithuania, The Netherlands, New Zealand, Northern Ireland, Norway, Slovak Republic, and Sweden. The percentage of those countries with strong input control mechanisms is slightly higher than that of the weak control ones. Specifically, 25 out of 43 countries (58 %) exercise strong input control over the teachers, i.e., their curriculum prescribes more than goals and objectives. As data analysis reveals, in most cases, national curricula in 16 countries also impose control in the remaining three areas. In addition to the goals and objectives, their national curricula prescribe instructional processes or methods, materials (e.g., textbooks, instructional materials), and assessment methods/activities. These countries are Azerbaijan, Bahrein, Chinese Taipei, Hong Kong, Iran, Kazakhstan, Macao, Morocco, Oman, Portugal, Qatar, Saudi Arabia, Singapore, Slovenia, Trinidad and Tobago, and the United Arab Emirates. The remaining ten countries with one or two areas prescribed by the national curricula in addition to the goals and objectives are Austria, Bulgaria, Canada, Chile, Georgia, Israel, Latvia, Russian Federation, and Spain.

Compared to the input control forms, a higher share of educational systems (70 %) makes a preference in favor of strong output control mechanisms, i.e., they use national and regional examinations to evaluate the implementation of the reading curriculum solely or in addition to other evaluation mechanisms (such as research programs, school self-evaluation, and visits by inspectors).

The list of these countries is as follows: Australia, Austria, Azerbaijan, Belgium (Flemish), Bulgaria, Canada, Chile, Chinese Taipei, Denmark, France, Georgia,

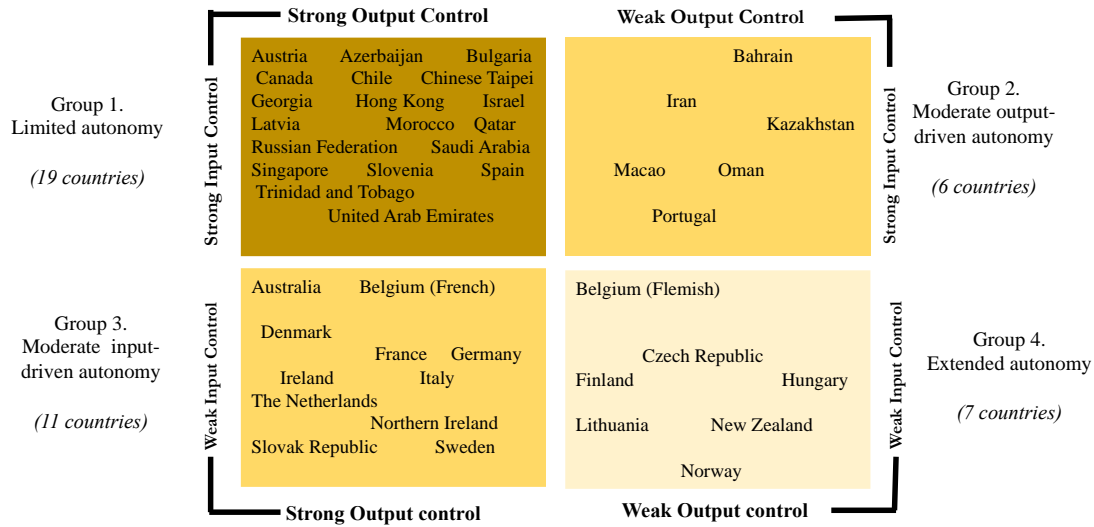


Germany, Hong Kong, Ireland, Israel, Italy, Latvia, Morocco, The Netherlands, Northern Ireland, Qatar, Russian Federation, Saudi Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Trinidad and Tobago, United Arab Emirates. The remaining 13 countries (30 %) exercise relatively soft control, i.e., they use various mechanisms (i.e., visits by inspectors, research programs, school self-evaluation), but national or regional examinations: Bahrain, Belgium (Flemish), Czech Republic, Finland, Hungary, Iran, Kazakhstan, Lithuania, Macao, New Zealand, Norway, Oman, and Portugal.

Based on the analysis of the countries by the types of input and output control, I categorized them into four groups by freedom from control, i.e., autonomy, incorporating both freedom from input and output control: (1) Limited autonomy (i.e., countries with strong input and strong output control mechanisms); (2) Moderate output-driven autonomy (i.e., countries with strong input and weak output control mechanisms); (3) Moderate input-driven autonomy (i.e., weak input and strong output), (4) Extended autonomy (i.e., weak input and weak output control mechanisms) (See Figure 2).

Figure 2.

*Distribution of Countries by Teacher Autonomy*



As shown in Table 7, the countries with limited teacher autonomy (group 1), i.e., with strong input and output control in the curriculum implementation domain, make the largest group with 19 countries. In addition to the size, it stands out as the most diverse group. Geographically, it includes countries from Western Europe (Austria, Spain), Eastern Europe (Slovenia, Georgia, Latvia, Bulgaria, Azerbaijan, Russian Federation), Middle East and North Africa (Israel, Qatar, Saudi Arabia, Morocco, United Arab Emirates), Latin America and the Caribbean (Chile, Trinidad and Tobago), North America (Canada), and East Asia (Chinese Taipei, Hong Kong, Singapore). There are significant variations in economic development as well. While the majority of the countries (14 out of 20) are high-income economies, there are upper-middle income (Azerbaijan, Bulgaria, Chinese Taipei, Russian Federation) and lower-middle income (Georgia, Morocco) economies represented as well. Countries with limited freedom from control also vary by the level of democratic development: three countries are considered

full democracies (Austria, Canada, and Spain), eight countries are regarded as flawed democracies (Bulgaria, Chile, Hong Kong, Israel, Latvia, Singapore, Slovenia, Trinidad and Tobago), two are defined as hybrid regimes (Georgia and Morocco), and six are considered as authoritarian regimes (Azerbaijan, Chinese Taipei, Qatar, Russian Federation, Saudi Arabia, United Arab Emirates).

The group of countries of moderate output-driven autonomy (Group 2), i.e., with strong input and weak output control in the curriculum domain, makes the smallest group with only six countries included (see Table 7). Geographically, it represents countries from the Middle East (Bahrain, Oman, Iran), Central Asia (Kazakhstan), Western Europe (Portugal), and East Asia (Macao). According to the 2016 classification by the World Bank, Iran and Kazakhstan are upper-middle-income countries, with the rest qualifying as high-income ones. Interestingly, by democratic development, all five countries, except Portugal (flawed democracy), belong to authoritarian regimes.

The countries with moderate input-driven autonomy (Group 3), i.e., weak input and strong output control mechanisms, make the second largest group among the four (see Table 7). Despite its large size, it is quite homogenous from the perspective of geographic representation and economic and democratic development. All countries represent Western and Eastern Europe (Belgium (French), Denmark, France, Germany, Ireland, Italy, the Netherlands, Northern Ireland, and Sweden)) and Eastern Europe (Slovak Republic), with the addition of Australia from the Pacific. In terms of economic development, all of them are high-income economies. Finally, all the countries have advanced democratic development and qualify as full or flawed democracies.

The countries with extended autonomy (Group 4), i.e., weak input and weak output control mechanisms, make another homogenous cluster (See Table 7). Geographically, it resembles group 3: Western Europe (Belgium (Flemish), Finland, Norway), Eastern Europe (Hungary, Czech Republic, Lithuania), and the Pacific (New Zealand). All the countries are advanced in terms of both economic and democratic development. By 2016, when PIRLS 2016 was administered, all qualified as high-income countries. By the Democracy Index, they ranked as full democracies (Finland, New Zealand, Norway) or flawed democracies (Belgium, Czech Republic, Hungary, Lithuania).

Table 6.

*PIRLS 2016 Countries: Scores and Types of Input and Output Control*

Country	Input Control Type	1. Input Control				2. Output Control				
		1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4	
1 Australia	Weak	1	0	0	0	Strong	0	0	1	1
2 Austria	Strong	1	0	0	1	Strong	1	0	1	1
3 Azerbaijan	Strong	1	1	1	1	Strong	0	1	1	1
4 Bahrain	Strong	1	1	1	1	Weak	1	0	0	0
5 Belgium (Flemish)	Weak	1	0	0	0	Weak	1	1	1	0
6 Belgium (French)	Weak	1	0	0	0	Strong	1	0	0	1
7 Bulgaria	Strong	1	1	0	0	Strong	1	1	1	1
8 Canada	Strong	1	0	0	1	Strong	0	0	1	1
9 Chile	Strong	1	1	0	0	Strong	1	1	0	1
10 Chinese Taipei	Strong	1	1	1	1	Strong	0	0	1	1
11 Czech Republic	Weak	1	0	0	0	Weak	1	0	1	0
12 Denmark	Weak	1	0	0	0	Strong	0	1	1	1
13 Finland	Weak	1	0	0	0	Weak	0	1	1	0
14 France	Weak	1	0	0	0	Strong	1	0	0	1
15 Georgia	Strong	1	0	1	0	Strong	0	0	0	1
16 Germany	Weak	1	0	0	0	Strong	1	0	1	1
17 Hong Kong, SAR	Strong	1	1	1	1	Strong	1	1	1	1
18 Hungary	Weak	1	0	0	0	Weak	1	0	1	0
19 Iran	Strong	1	1	1	1	Weak	1	1	0	0
20 Ireland	Weak	1	0	0	0	Strong	1	1	1	1
21 Israel	Strong	1	1	1	0	Strong	1	0	1	1
22 Italy	Weak	1	0	0	0	Strong	1	1	1	1
23 Kazakhstan	Strong	1	1	1	1	Weak	1	1	1	0

24	Latvia	Strong	1	0	0	1	Strong	1	0	1	1
25	Lithuania	Weak	1	0	0	0	Weak	0	0	0	0
26	Macao	Strong	1	1	1	1	Weak	1	0	1	0
27	Morocco	Strong	1	1	1	1	Strong	1	1	0	1
28	The Netherlands	Weak	0	0	0	0	Strong	1	1	1	1
29	New Zealand	Weak	1	0	0	0	Weak	1	1	1	0
30	Northern Ireland	Weak	1	0	0	0	Strong	1	1	1	1
31	Norway	Weak	1	0	0	0	Weak	0	1	1	0
32	Oman	Strong	1	1	1	1	Weak	1	1	1	0
33	Portugal	Strong	1	1	1	1	Weak	1	0	1	0
34	Qatar	Strong	1	1	1	1	Strong	1	1	1	1
35	Russian Federation	Strong	1	0	1	0	Strong	0	0	1	1
36	Saudi Arabia	Strong	1	1	1	1	Strong	1	1	1	1
37	Singapore	Strong	1	1	1	1	Strong	0	1	1	1
38	Slovak Republic	Weak	1	0	0	0	Strong	1	0	1	1
39	Slovenia	Strong	1	1	1	1	Strong	0	1	1	1
40	Spain	Strong	1	1	0	1	Strong	0	0	1	1
41	Sweden	Weak	1	0	0	0	Strong	1	1	1	1
42	Trinidad and Tobago	Strong	1	1	1	1	Strong	1	0	0	1
43	United Arab Emirates	Strong	1	1	1	1	Strong	1	0	0	1

Input Control:

- 1.1. The language/reading curriculum prescribes goals and objectives
- 1.2. The language/reading curriculum prescribes instructional processes or methods
- 1.3. The language/reading curriculum prescribes materials (e.g., textbooks, instructional materials)
- 1.4. The language/reading curriculum prescribes assessment methods/activities

Output Control:

- 2.1. The language/reading curriculum implementation is evaluated by visits by inspectors
- 2.2. The language/reading curriculum implementation is evaluated by research programs
- 2.3. The language/reading curriculum implementation is evaluated by school self-evaluation
- 2.4. The language/reading curriculum implementation is evaluated by national or regional examinations

Explanation of scores:

1 – Yes, 0 - No

### ***Summary of the findings***

Analysis of the distribution of countries by the type of autonomy, defined as freedom from control, reveals that high-income countries with advanced democratic development from Europe and the Pacific region and Oceania mostly institutionalize weak input national-level control mechanisms for curriculum implementation paired with weak or strong output control mechanisms. Therefore, teachers in these countries enjoy extended or moderate input-driven freedom from the national level control in curriculum implementation. Governments with authoritarian and hybrid regimes, irrespective of

economic development, tend to implement strong input control mechanisms combined with strong output control (in the majority of cases) or weak output control forms. As a result, teachers are entitled to either limited freedom or output-oriented freedom from control in curriculum implementation.

Table 7.

*Countries by Autonomy, Economic and Democratic Development*

<b>Country</b>	<b>Input Control</b>	<b>Output Control</b>	<b>Economic development</b>	<b>Democratic development</b>
<b>Group 1. Limited autonomy</b>				
Austria	Strong	Strong	High income	Full democracy
Azerbaijan	Strong	Strong	Upper middle income	Authoritarian
Bulgaria	Strong	Strong	Upper middle income	Flawed democracy
Canada	Strong	Strong	High income	Full democracy
Chile	Strong	Strong	High income	Flawed democracy
Chinese Taipei	Strong	Strong	Upper middle income	Authoritarian**
Georgia	Strong	Strong	Lower middle income	Hybrid Regime
Hong Kong, SAR	Strong	Strong	High income	Flawed democracy
Israel	Strong	Strong	High income	Flawed democracy
Latvia	Strong	Strong	High income	Flawed democracy
Morocco	Strong	Strong	Lower middle income	Hybrid Regime
Qatar	Strong	Strong	High income	Authoritarian
Russian Federation	Strong	Strong	Upper middle income	Authoritarian
Saudi Arabia	Strong	Strong	High income	Authoritarian
Singapore	Strong	Strong	High income	Flawed democracy
Slovenia	Strong	Strong	High income	Flawed democracy
Spain	Strong	Strong	High income	Full democracy
Trinidad and Tobago	Strong	Strong	High income	Flawed democracy
United Arab Emirates	Strong	Strong	High income	Authoritarian
<b>Group 2. Moderate output-driven autonomy</b>				
Bahrain	Strong	Weak	High income	Authoritarian
Iran	Strong	Weak	Upper middle income	Authoritarian
Kazakhstan	Strong	Weak	Upper middle income	Authoritarian
Macao	Strong	Weak	High income	Authoritarian**
Oman	Strong	Weak	High income	Authoritarian
Portugal	Strong	Weak	High income	Flawed democracy
<b>Group 3. Moderate input-driven autonomy</b>				
Australia	Weak	Strong	High income	Full democracy
Belgium (French)	Weak	Strong	High income	Flawed democracy
Denmark	Weak	Strong	High income	Full democracy

<b>Country</b>	<b>Input Control</b>	<b>Output Control</b>	<b>Economic development</b>	<b>Democratic development</b>
France	Weak	Strong	High income	Flawed democracy
Germany	Weak	Strong	High income	Full democracy
Ireland	Weak	Strong	High income	Full democracy
Italy	Weak	Strong	High income	Flawed democracy
The Netherlands	Weak	Strong	High income	Full democracy
Northern Ireland	Weak	Strong	High income*	Full democracy*
Slovak Republic	Weak	Strong	High income	Flawed democracy
Sweden	Weak	Strong	High income	Full democracy
<b>Group 4. Extended autonomy</b>				
Belgium (Flemish)	Weak	Weak	High income	Flawed democracy
Czech Republic	Weak	Weak	High income	Flawed democracy
Finland	Weak	Weak	High income	Full democracy
Hungary	Weak	Weak	High income	Flawed democracy
Lithuania	Weak	Weak	High income	Flawed democracy
New Zealand	Weak	Weak	High income	Full democracy
Norway	Weak	Weak	High income	Full democracy
* United Kingdom data	** China data			

### **Differences in Student Achievement and Achievement Gap by Type of Autonomy**

To explore whether student achievement in reading, as well as the achievement gap, varied across four groups of educational systems by teacher autonomy (limited, moderate output-driven, moderate input-driven, and extended), I conducted a one-way analysis of variance (ANOVA), two-way ANOVA and an independent Samples t-Test. I also examined variation by country's level of economic and democratic development, achievement quartiles, student socioeconomic background, and share of economically disadvantaged students at school. When addressing this research question, I did not control for student, teacher, school, or country-level variables. Therefore, the differences in student achievement and achievement gap cannot be solely attributed to the type of teacher autonomy or any other factors I have included in the analysis.

I used student average score in reading, a continuous measure, as a dependent variable across all analyses. In addition, I used the following measures as predictor variables in respective analysis:

- Country's economic development level: (1) lower-middle income, (2) upper-middle income, (3) high-income.
- Country's democratic development level: (1) authoritarian regime, (2) hybrid regime, (3) flawed democracy, (4) full democracy.
- Share of economically disadvantaged students at school: (1) 0-10 %, (2) 11 – 25 %, (3) 26 – 50 %, (4) More than 50%.
- Student socioeconomic background (home resources): (1) few resources, (2) some resources, (3) many resources.
- Achievement quartiles: (1) bottom quartile, (2) top quartile.

### ***Differences in Student Achievement***

I performed a one-way ANOVA to understand whether student achievement in reading differed across the groups of countries with different forms of teacher autonomy in curriculum implementation. The test revealed a statistically significant difference in average student reading achievement scores across all four groups ( $p < 0.001$ ) with an effect size of medium magnitude ( $\eta^2 = 0.104$ ). Also, Bonferroni post hoc test revealed a statistically significant difference between all groups.

As demonstrated in Table 8, Students in countries with moderate output-driven autonomy showed far lower average reading scores ( $M_2 = 453.54$ ,  $SD_2 = 101.12$ ) than their peers in all other groups of countries. Reading scores were the highest, with far smaller variation among the students from the countries with extended autonomy ( $M_4 =$



551.49,  $SD_4 = 66.07$ ). Overall, on average, students in countries with weak output control, either in a combination of strong or weak output control mechanisms (i.e., with moderate input-driven and extended autonomy), performed better. Groups of countries with strong input control mechanisms (i.e., with limited and moderate output-driven autonomy) had the lowest average achievement compared to those with weak input control (i.e., extended or moderate input-driven autonomy). However, students' average reading score was significantly higher when strong input control measures were paired with strong output control measures, i.e., when teachers had limited autonomy ( $M_2 = 453.54$ ,  $SD_2 = 66.07$  vs.  $M_1 = 521.47$ ;  $SD_1 = 106.47$ ).

I also examined the distribution of average reading scores by countries' economic development level. As mentioned, Group 3 (moderate input-driven autonomy) and Group 4 (extended autonomy) comprise only high-income economies. Therefore, I examined the differences within Group 1 (limited autonomy) and Group 2 (moderate output-driven autonomy). Table 8 depicts the results of one-way ANOVA for Group 1 and those of independent samples t-Test for Group 2. As shown, in Group 1 (limited autonomy), students in lower-middle-income countries scored the lowest ( $M = 361.26$ ,  $SD = 102.53$ ), and those in upper-middle-income countries – the highest ( $M = 574.87$ ,  $SD = 66.59$ ). Students from the high-income economies demonstrated middle-level performance ( $M = 513.15$ ,  $SD = 87.51$ ) compared to the peers from the other two groups ( $F(2, 3187206) = 1166997.49$ ,  $\eta^2 = 0.423$ ,  $p < 0.001$ ). According to the Bonferroni post hoc test results, differences were statistically significant between all three groups. I found a significant difference in the average reading scores between high-income economies ( $M = 494.60$ ,

$SD = 90.48$ ) and upper-middle-income countries ( $M = 448.99$ ,  $SD = 101.21$ ) in Group 2 (moderate output-driven autonomy) as well ( $t(1888460.03) = -1335.077$ ,  $p < 0.001$ ).

In addition to economic development, I examined the differences in average reading scores by country's democratic development within each group by teacher autonomy. One-way ANOVA demonstrated a statistically significant difference in the average reading score among the countries classified as authoritarian regimes ( $M = 552.17$ ,  $SD = 89.79$ ), hybrid regimes ( $M = 361.27$ ,  $SD = 102.53$ ), flawed democracies ( $M=532.26$ ,  $SD=80.28$ ), and full democracy ( $M = 538.85$ ,  $SD = 63.90$ ) in the group of countries with limited autonomy ( $F(3, 3187205) = 585131.716$ ,  $\eta^2 = 0.355$ ,  $p < 0.001$ )). According to the Bonferroni post hoc test results, differences were statistically significant between all four groups. An independent samples t-Test revealed a small but significant difference in the average reading score in flawed democracies ( $M = 528.62$ ,  $SD = 66.21$ ) and full democracies ( $M = 556.47$ ,  $SD = 66.51$ ) within the group of countries with moderate input-driven autonomy (as well ( $t(1435197.59) = -266.777$ ,  $p < 0.001$ )). Similarly, the difference in the average reading score between these two groups by democratic development was found to be statistically significant ( $t(232204.219) = -69.083$ ,  $p < 0.001$ ) in the group of countries with extended autonomy: flawed democracies ( $M = 546.00$ ,  $SD = 65.19$ ) and full democracies ( $M = 562.37$ ,  $SD = 65.19$ ). Finally, in the group of countries with moderate output-driven freedom from control, I found the average reading score of students in hybrid regimes ( $M = 448.99$ ,  $SD = 101.21$ ) far lower than that of those in flawed democracies ( $M = 529.74$ ,  $SD = 61.22$ ). Similar to the other groups, this difference was statistically significant ( $t(120685.669)=-360.043$ ,  $p < 0.001$ ).

Table 8.

*Student Achievement by Teacher Autonomy*

	Country groups by teacher autonomy			
	Group 1: Limited	Group 2: Moderate Output-driven	Group 3: Moderate Input-driven	Group 4: Extended
Average reading score	521.47 (106.47)	453.54 (101.12)	539.72 (67.72)	551.49 (66.07)
	F(6625021)=255695.858 $p < 0.001$			
Average reading score				
Lower-middle-income	361.26 (102.53)	448.99 (101.21)		
Upper-middle-income	574.87 (66.59)	494.60 (90.48)		
High-income	513.15 (87.51)		539.72 (67.72)	551.49 (66.07)
	$F(2, 3187206) =$ 1166997.49 $\eta^2 = .423$ $p < 0.001$	$t(1888460.03) =$ 1335.077 $p < 0.001$		
Average reading score				
Authoritarian regimes	552.17 (89.79)	448.53 (101.21)	-----	-----
Hybrid regimes	361.27 (102.53)	-----	-----	-----
Flawed democracy	532.26 (80.28)	529.74 (61.22)	528.62 (66.21)	546.00 (65.83)
Full democracy	538.85 (63.90)	-----	556.47 (66.51)	562.37 (65.19)
	$F(3, 3187205) =$ 585131.716 $\eta^2 = 0.355$ $p < 0.001$	$t(120685.669) =$ -360.043 $p < 0.001$	$t(1435197.59) =$ -266.777 $p < 0.001$	$t(232204.219) =$ 69.083 $p < 0.001$

*Differences in Student Achievement Gap*

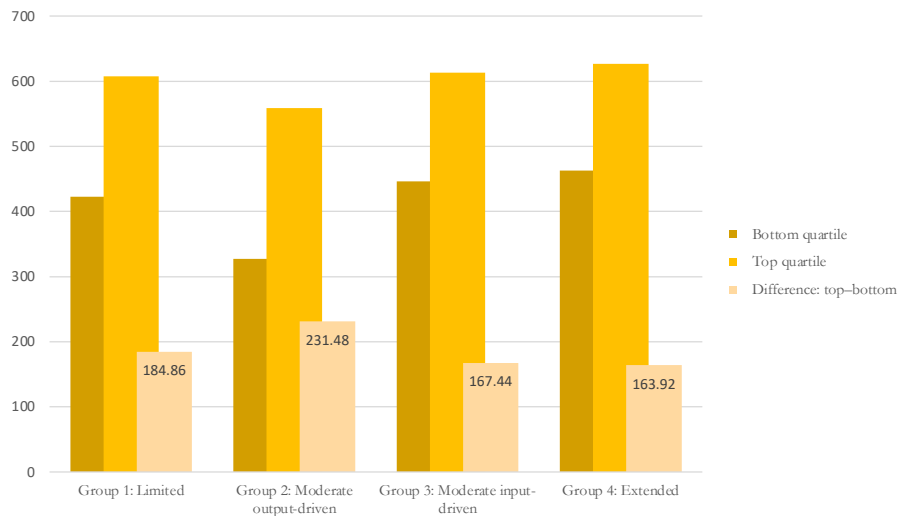
To explore how teacher autonomy is associated with student achievement in reading, I examined the distribution by the top (75<sup>th</sup> percentile) and bottom (25<sup>th</sup> percentile) quartiles of reading scores in the PIRLS 2016 distribution. In addition, I explored differences in average student achievement by student socioeconomic background (i.e., home resources for learning) and school SES (the share of economically disadvantaged students at schools) across four groups by teacher autonomy.

First, I analyzed the distribution of average student scores in reading by students in the top and bottom quartiles. Independent samples t-Test revealed a statistically

significant difference in the average scores between students in the top and bottom quartiles across all four groups of countries ( $p < 0.001$ ). As Table 9 and Figure 3 show, the difference between the average scores of students in the top and bottom quartiles is the highest in the group of countries with moderate output-driven freedom from control (231.48 points), followed by the group of countries with limited freedom from control (184.86 points). The achievement gap was found to be the smallest in the groups of countries with moderate input-driven and extended freedom from control (167.44 and 163.92 points, respectively).

Figure 3.

*Achievement Gap by Achievement Quartiles*



Next, I compared the distribution of average student achievement in reading by the student's socioeconomic background, i.e., the amount of available resources at home (categorized as few, some, and many) and by type of teacher autonomy. Table 9 and Figure 4 depict the results of the one-way ANOVA test I conducted for each of the four groups of countries. As these results indicate, the differences in average scores within

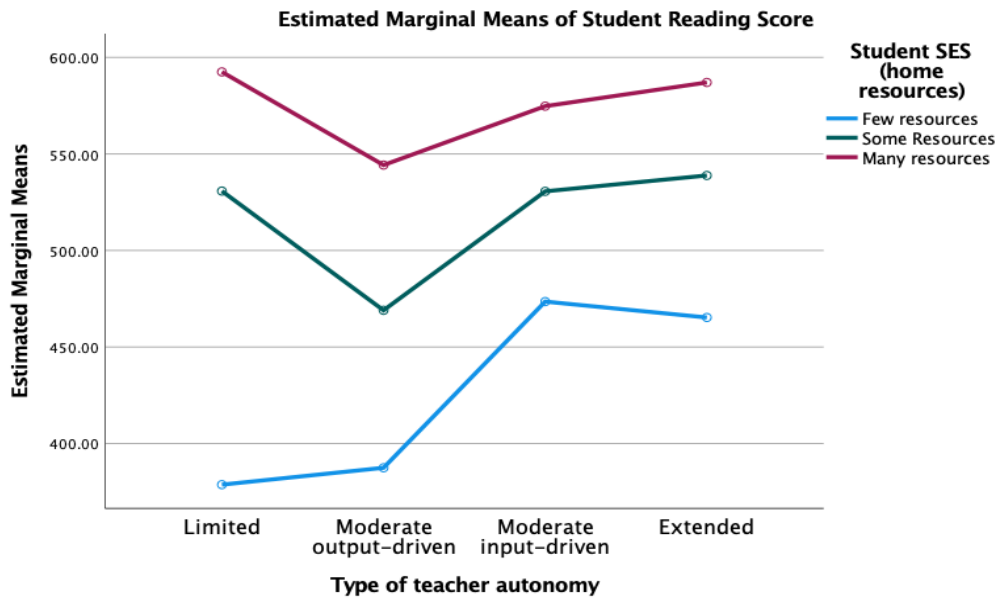
each group between students with few, some, and many resources were found to be statistically significant ( $p < 0.001$ ). The multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average student achievement between all SES categories of students within all four groups of countries by autonomy. The effect size, measured by eta-squared, was large in all groups (limited:  $\eta^2 = 0.281$ ; moderate output-driven:  $\eta^2 = 0.158$ ; extended:  $\eta^2 = 0.153$ ), except for moderate input-driven ( $\eta^2 = 0.103$ ). To analyze whether the type of freedom from control and the SES of a student are jointly associated with a student's achievement score reading, I also conducted a two-way ANOVA test. The analysis yielded a statistically significant interaction between the type of freedom from control and SES in how they jointly are associated with student reading scores ( $F(6, 6626475) = 23900.100, p < 0.001$ ).

As evident from Figure 4 and Table 9, countries with limited and moderate output-driven freedom from control have the largest gap in achievement by student socioeconomic background. These differences are the smallest in the countries with extended and moderate input-driven freedom from control. As an example, the difference in scores between students with many and few resources range from 213.79 points in Group 1 (extended autonomy) to 101.15 points in Group 3 (moderate input-driven autonomy). It should be noted that average scores for students with many resources are quite similar across the groups with different types of teacher autonomy. For example, high SES students in countries with limited autonomy perform as high as their counterparts in countries with extended one ( $M_1 = 592.53, SD_1 = 62.44; M_4 = 587.06, SD_4 = 56.02$ ). At the same time, as expected from the high achievement gap, students with few resources performed far below the students with the similar SES background in the

other two groups ( $M_1 = 378.74$ ,  $SD_1 = 111.35$ ;  $M_2 = 387.45$ ,  $SD_2 = 101.06$  vs.  $M_3 = 473.64$ ,  $SD_3 = 72.23$ ;  $M_4 = 465.33$ ,  $SD_4 = 63.19$ ).

Figure 4.

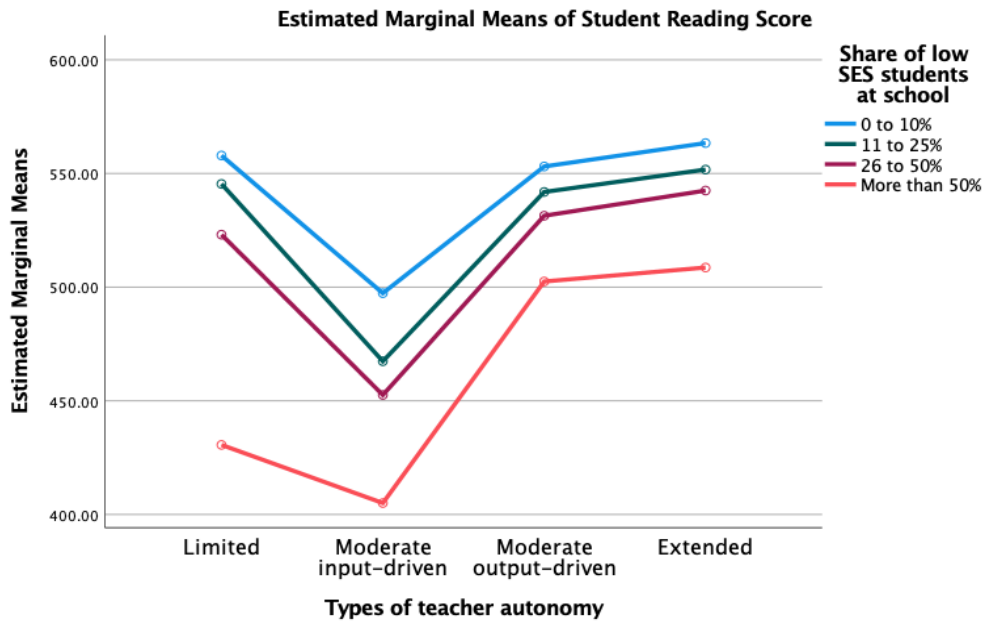
*Achievement Gap by Student Socioeconomic Background*



A similar investigation of the achievement gap by the share of economically disadvantaged students at schools showed that countries with limited autonomy have the highest discrepancy in student achievement. Namely, as shown in Table 9 and Figure 5, the difference between the average score of students from the lowest (0-10 % of disadvantaged students) and highest (with more than 50 % of disadvantaged students) share of low SES students is as high as 127.31 points. Countries with moderate output-driven autonomy follow them close with 92.32 points. An achievement gap is considerably low in countries with moderate input-driven and extended freedom from control (difference of 50.55 and 54.72 points, respectively).

Figure 5.

*Achievement Gap by the Share of Disadvantaged Students*



The differences in average scores within each group between students who attend schools with a different share of economically disadvantaged students were found to be statistically significant ( $p < 0.001$ ). The multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average student achievement between all categories of students with varying school SES within all four groups of countries by freedom from control. One-way ANOVA test revealed a large effect size in the groups of countries with limited and moderate output-driven freedom from control ( $\eta^2 = 0.210$  and  $\eta^2 = 0.124$ , respectively). In the remaining two countries with the lower achievement gap, the effect of school composition on student achievement was found to be medium with the eta-squared values of 0.053 and 0.058. I also conducted a two-way ANOVA test that showed that there is a statistically significant interaction between the type of freedom from control and school composition by SES in how they are associated with student reading scores ( $F(9, 6626471) = 14423.109, p < 0.001$ ).

Table 9.

*Achievement Gap by Teacher Autonomy*

	Country groups by teacher autonomy			
	Group 1: Limited	Group 2: Moderate output-driven	Group 3: Moderate input-driven	Group 4: Extended
<b>The achievement gap in reading by quartiles</b>				
Bottom quartile	422.82 (102.50)	327.54 (83.04)	445.91 (43.41)	462.71 (42.23)
Top quartile	607.68 (63.52)	559.02 (39.33)	13.35 (31.47)	626.63 (29.41)
Difference: bottom-top	184.86	231.48	167.44	163.92
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
<b>The achievement gap in reading by student socioeconomic background (home resources)</b>				
Few resources	378.74 (111.35)	387.45 (101.06)	473.64 (72.23)	465.33 (63.19)
Some resources	530.82 (90.95)	469.04 (91.79)	530.67 (65.19)	538.88 (62.72)
Many resources	592.53 (62.44)	544.29 (60.35)	574.79 (59.80)	587.06 (56.02)
Difference (many-few)	213.79	156.84	101.15	121.73
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	2, 3187206= 623755.699	2, 1408689= 132447.637	2, 1687105= 96648.188	2, 342011= 30904.149
Eta-squared ( $\eta^2$ )	0.281	0.158	0.103	0.153
<b>The achievement gap in reading by the share of economically disadvantaged students at school</b>				
0 – 10 %	557.97 (80.83)	497.39 (79.68)	553.12 (64.78)	563.37 (61.98)
11 – 25 %	545.45 (86.49)	467.51 (96.91)	541.90 (68.28)	551.80 (62.71)
26 – 50 %	523.14 (96.19)	452.59 (101.86)	531.45 (68.28)	542.51 (65.60)
More than 50 %	430.66 (120.42)	405.07 (99.89)	502.57 (71.19)	508.65 (75.44)
Difference (0-10 % - more than 50 %)	127.31	92.32	50.55	54.72
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	3, 3187205= 282632.674	3, 1408688 = 66194.058	3, 1687107= 31206.218	3, 342010= 6955.536
Eta-squared ( $\eta^2$ )	0.210	0.124	0.053	0.058

***Summary of the findings***

The results of independent sample t-Test, one-way ANOVA, and two-way ANOVA revealed significant variations in student achievement across countries with different forms of teacher autonomy in reading curriculum implementation. Overall average student achievement is the highest in countries with extended and moderate input-oriented autonomy. Students in the remaining two groups fall behind, with those in the countries with moderate output-driven autonomy performing the worst. In addition,



the breakdown of average student achievement by country's economic development shows that within the group with limited autonomy, the upper-middle-income ones have the highest average student performance. In contrast, the high-income countries exceeded the upper-middle-income ones in the systems with moderate output-driven autonomy. There are variations in the level of democratic development as well. In the systems with limited autonomy, on average, students from the countries with authoritarian regimes demonstrate the highest achievement, while the ones from the hybrid regimes - are the lowest. Across the systems with moderate output-driven autonomy, students in authoritarian regimes fall far behind those in flawed democracies. Finally, in the systems with moderate input-driven and extended teacher autonomy, students from full democracies slightly outperform their peers from flawed democracies.

The findings also indicate variation in the achievement gap. Two groups of countries with the lowest student achievement (limited and moderate output-driven) also show considerable achievement gaps. Advantaged students, i.e., those whose families have many resources and who study at schools with a small share of economically disadvantaged students, significantly outperform their disadvantaged peers. This magnitude of achievement discrepancy is particularly alarming for counties with limited autonomy. Findings reveal that national systems with extended and moderate input-driven autonomy have significantly and consistently lower achievement gaps.

### **Differences in Teachers' Use of Classroom Practices Type of Autonomy**

To understand whether teachers' use of selected classroom practices varied across four groups of systems by teacher autonomy (limited, moderate output-driven, moderate input-driven, and extended), I conducted a one-way analysis of variance (ANOVA), two-

way ANOVA, and chi-square test. In addition to the general overview, I examined variations in the use of practices by the share of low SES students at school and teachers' perceptions about the limitations related to student attributes. When addressing this research question, I did not control for student, teacher, school, or country-level variables. Therefore, the differences in using these classroom practices cannot be solely attributed to the type of teacher autonomy or any additional factors I have included in the analysis.

I used classroom-practice-related variables as dependent variables across all tests. Among them, reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies were continuous measures. Individualized instruction was a categorical variable with two levels: (a) never and sometimes; and (b) often and always. For respective analysis, I used the following independent variables:

- Teacher autonomy: limited, moderate output-driven, moderate input-driven, extended.
- Share of economically disadvantaged students at school: (1) 0-10 %, (2) 11-25 %, (3) 26-50 %, (4) More than 50%.
- Instruction limited by student attributes (1) very little, (2) some, (3) a lot.

### ***General Overview of the Teachers' Use of Classroom Practices by Teacher Autonomy***

First, I conducted a one-way analysis of variance (ANOVA) to detect the differences in the frequency of the use of reading comprehension, content differentiation and autonomy-supportive strategies across the systems with different types of autonomy. Due to the categorical nature of the individualized instruction variable, I used a chi-

square test to examine the similar relationship between teacher autonomy and the use of individualized instruction.

**Use of reading comprehension strategies.** As Table 10 displays, average scores indicating the frequency of teachers' use of strategies for reading comprehension varied across the systems with a different forms of teacher autonomy. On average, teachers in countries with limited and moderate output-driven autonomy used these strategies most frequently ( $M_1 = 30.3$ ,  $SD_1 = 4.7$ ;  $M_2 = 29.6$ ,  $SD_2 = 5.3$ ). The indicators of the frequency of the use of strategies were significantly lower in the systems with moderate input-driven and extended autonomy ( $M_3 = 27.0$ ,  $SD_3 = 5.1$ ;  $M_4 = 27.3$ ,  $SD_4 = 5.1$ ). According to the results of the one-way ANOVA test, these differences were statistically significant ( $p < 0.001$ ) with an effect size of medium magnitude as measured by Eta squared ( $\eta^2 = 0.079$ ).

**Use of reading content differentiation.** Table 10 shows that, on average, content differentiation by level and interest was most frequently used by the teachers in the systems with limited or moderate input-driven autonomy ( $M_1 = 6.0$ ,  $SD_1 = 1.5$ ;  $M_3 = 5.8$ ,  $SD_3 = 1.4$ ). Compared to the teachers from these two groups of countries, those from input-driven autonomy differentiated reading materials less frequently ( $M_2 = 5.5$ ,  $SD_2 = 1.6$ ). However, on average, teachers with extended autonomy reported the least frequent use ( $M_4 = 5.4$ ,  $SD_4 = 1.4$ ). According to the results of the one-way ANOVA test, these differences were statistically significant ( $p < 0.001$ ) with an effect size of small magnitude ( $\eta^2 = 0.020$ ).

**Use of autonomy-supportive strategies.** Similar to the findings on content differentiation, teachers from the group of extended autonomy reported the least frequent

use of autonomy-supportive strategies ( $M_4 = 20.6, SD_4 = 3.3$ ). The same values for the group with moderate input-driven autonomy were also one of the lowest ( $M_3 = 20.8, SD_3 = 3.3$ ). The highest frequency of using these strategies was shown in the systems with moderate output-driven autonomy ( $M_2 = 23.3, SD_2 = 3.3$ ), followed by those with limited one ( $M_1 = 21.7, SD_1 = 3.3$ ). One-way ANOVA test revealed that these differences were statistically significant ( $p < 0.001$ ) with an effect size of medium magnitude ( $\eta^2 = 0.072$ ).

Table 10.

*Classroom Practices by Teacher Autonomy*

Classroom Practices	Country groups by teacher autonomy				F Statistics	p-value	Eta-Squared
	Group 1: Limited	Group 2: Moderate Output-driven	Group 3: Moderate Input-driven	Group 4: Extended			
Reading comprehension strategies	30.3 (4.7)	29.6 (5.3)	27.0 (5.1)	27.3 (5.1)	3,6620521=188938.472	<0.001	0.079
Reading content differentiation	6.0 (1.5)	5.5 (1.6)	5.8 (1.4)	5.4 (1.4)	3,6620521=45459.176	<0.001	0.020
Autonomy-supportive strategies	21.7 (3.3)	23.3 (3.3)	20.8 (3.3)	20.6 (3.3)	3,6620521=171565.284	<0.001	0.072

**Use of individualized instruction.** The chi-square test of association revealed a significant association between teacher autonomy and the use of individualized instruction ( $p < 0.001$ ). As Table 11 shows, countries with limited autonomy had the highest share of teachers who often or always use individualized instruction (50.4 %). This number was the smallest in groups with moderate input-driven and extended autonomy (28.8 % and 37.2 %, respectively).

Table 11.

*Individualized Instruction by Teacher Autonomy*

Country groups by teacher autonomy	Teachers Using Individualized Instruction (%)	
	Never or Sometimes	Often or Always
Group 1: Limited	49.6	50.4
Group 2: Moderate output-driven	56.0	44.0
Group 3: Moderate input-driven	71.2	28.8
Group 4: Extended	62.8	37.2

Overall, teachers from the countries with limited and output-driven autonomy made more frequent use of reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies compared to those from input-driven and extended autonomy. Also, a higher share of teachers from these first two groups of countries used individualized instruction.

### ***Teachers' Use of Classroom Practices by the Share of Economically Disadvantaged Students at School***

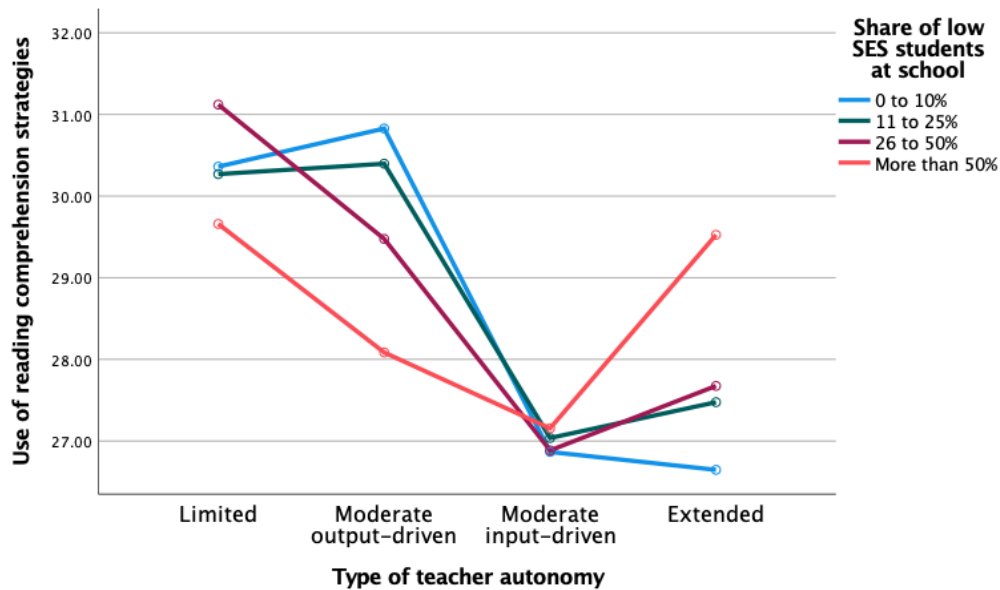
In addition to the overall distribution of the frequency of teacher use of four key practices across the systems with different forms of teacher autonomy, I examined how teachers' application of these practices varied by schools with a different share of economically disadvantaged students. For this purpose, I used one-way ANOVA and two-way ANOVA, as well as a chi-square test (to analyze the use of individualized instruction).

**Use of reading comprehension strategies.** One-way ANOVA test revealed statistically significant differences ( $p < 0.001$ ) in the average use of reading comprehension strategies between teachers who teach at schools with different share of low SES students across all four groups of countries. Except for one case (see Table 12), the multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of reading comprehension strategies between all categories of teachers with a varying share of low SES students within all four groups of countries

by autonomy. The effect size, as measured by Eta-squared, was found to be of small magnitude in three groups of countries (limited:  $\eta^2 = 0.009$ ; moderate input-driven:  $\eta^2 = 0.000$ ; extended:  $\eta^2 = 0.027$ ) except for the group with moderate output-driven autonomy with a medium effect ( $\eta^2 = 0.045$ ). In addition, a two-way ANOVA test detected a statistically significant interaction between the school's SES composition and type of teacher autonomy in how they jointly associated with teacher use of reading comprehension strategies ( $F(9, 6626471) = 7723.346, p < 0.001$ ).

Figure 6.

*Reading Comprehension Strategies by the Share of Disadvantaged Students*



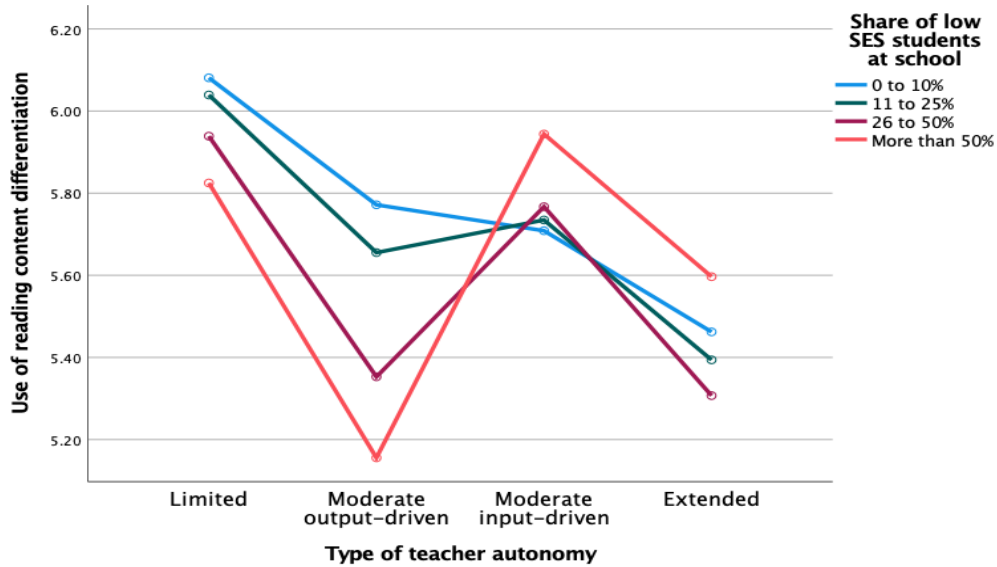
As depicted in Table 12 and Figure 6, in countries with moderate output-driven autonomy, the frequency of teacher use of reading comprehension strategies decreased as the percentage of low SES students increased (0-10 %:  $M_2=30.8, SD_2=4.6$ ; Vs. More than 50 %:  $M_2 = 28.1, SD_2 = 5.3$ ). Contrary to this, in countries with extended autonomy,

the frequency of teacher use of reading comprehension strategies increased along with the percentage of low SES students (0-10 %:  $M_4 = 26.7$ ,  $SD_4 = 5.1$  Vs. More than 50 %:  $M_4 = 29.5$ ,  $SD_4 = 4.9$ ). Similar to those in the countries with output-driven autonomy, those teachers with limited one, who teach at more disadvantaged schools, for the most part, used reading comprehension strategies more frequently (except for the teachers at schools with 26 to 50 % of low SES students, who reported the most frequent use of this practice ( $M_1 = 31.1$ ,  $SD_4 = 4.2$ )). Finally, the difference in the frequency of the use of reading comprehension strategies was almost similar across all the groups by the share of low SES students in the group of countries with moderate input-driven autonomy.

**Use of reading content differentiation.** One-way ANOVA test revealed statistically significant differences ( $p < 0.001$ ) in the average use of reading content differentiation between teachers who taught at schools with different percentages of low SES students across all four groups of countries. The multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of these strategies between all categories of teachers within all four groups of countries by teacher autonomy. A two-way ANOVA test detected a statistically significant interaction between the school's SES composition and type of autonomy in how they jointly associated with teacher use of content differentiation ( $F(9, 6626471) = 3689.752$ ,  $p < 0.001$ ).

Figure 7.

*Reading Content Differentiation by the Share of Disadvantaged Students*



As depicted in Table 12 and Figure 7, teacher use of content differentiation follows the same pattern as that of strategies for reading comprehension. In countries with limited and moderate output-driven autonomy, the frequency of teacher use of reading comprehension strategies decreased as the percentage of low SES students increased (0 -10 %:  $M_1 = 6.1$ ,  $SD_1 = 1.5$ ;  $M_2 = 5.7$ ,  $SD_2 = 1.6$  Vs. More than 50 %:  $M_1 = 5.8$ ,  $SD_1 = 1.6$ ;  $M_2 = 5.1$ ,  $SD_2 = 1.7$ ). Although, as shown, the magnitude of differences was larger for the teachers from the group of countries with moderate output-driven autonomy. In countries with moderate input-driven and extended ones, the differences in the frequency of teacher use of reading differentiation across these different types of schools were small, yet favored those schools that had low SES students as more than half of their student body (0 -10 %:  $M_3 = 5.7$ ,  $SD_3 = 1.3$ ;  $M_4 = 5.5$ ,  $SD_4 = 1.4$  Vs. More than 50 %:  $M_3 = 5.9$ ,  $SD_3 = 1.5$ ;  $M_4 = 5.6$ ,  $SD_4 = 1.4$ ). However, the effect size, as



measured by eta-squared, was small in all four groups of countries (limited:  $\eta^2 = 0.004$ ; moderate output-driven:  $\eta^2 = 0.024$ ; moderate input-driven:  $\eta^2 = 0.002$ , extended:  $\eta^2 = 0.003$ ).

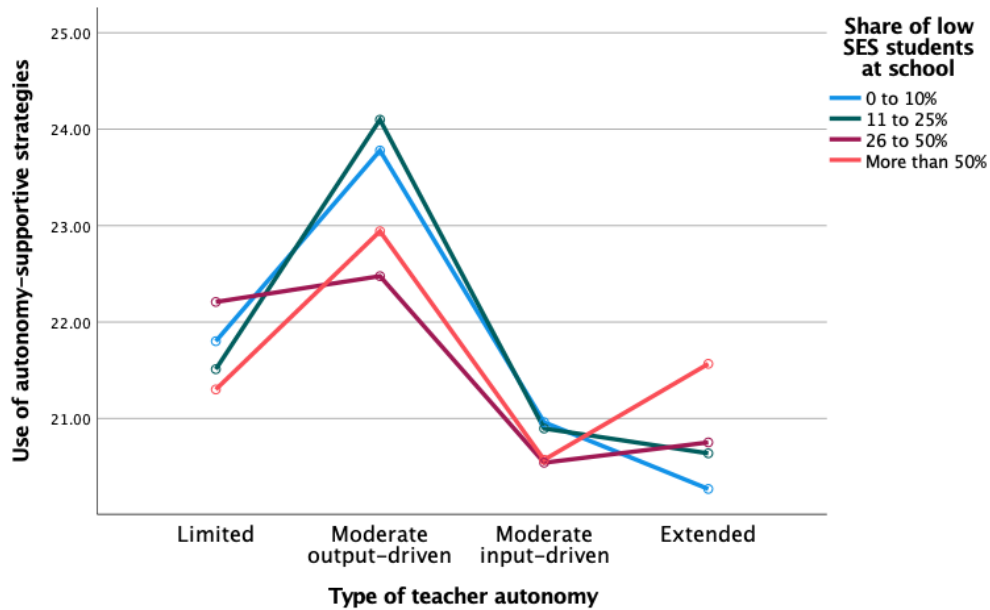
**Use of autonomy-supportive strategies.** I also examined the distribution of the use of autonomy-supportive strategies given the composition of schools by student socioeconomic background. The results of the one-way ANOVA test are shown in Table 12 and Figure 8. The differences in the use of strategies by teachers who taught at schools with different SES compositions were found to be statistically significant ( $p < 0.001$ ). The multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of reading comprehension strategies between all categories of teachers within all four groups of countries by autonomy. A two-way ANOVA test revealed a statistically significant interaction between the school's SES composition and type of autonomy on teacher use of autonomy-supportive strategies ( $F(9, 6626471) = 7580.013, p < 0.001$ )).

As Table 12 and Figure 8 demonstrate, compared to the use of other practices, there was less variation in the use of autonomy-supportive strategies among the teachers from schools with different share of low SES students. Yet, in the groups of countries with limited and moderate output-driven autonomy, teachers at schools with the lowest share of disadvantaged students on average used them more frequently than those at schools with the highest percentage of low SES students (0-10 %:  $M_1 = 21.8, SD_1 = 3.1$ ;  $M_2 = 23.8, SD_2 = 2.7$  Vs. More than 50 %:  $M_1 = 21.3, SD_1 = 3.4$ ;  $M_2 = 22.9, SD_2 = 3.5$ ). The group of schools with 26-50% of low SES students in the systems with limited

autonomy made the only exception from this pattern, as teachers from these schools reported the most frequent use of autonomy-supportive strategies ( $M_1 = 22.2$ ,  $SD_1 = 3.2$ ).

Figure 8.

*Autonomy-supportive Strategies by the Share of Disadvantaged Students*



The difference in the frequency of the use of these strategies was almost similar yet favoring advantaged students across all the groups by student SES composition in the group with moderate input-driven autonomy. As for the fourth group of countries with extended autonomy, as the percentage of low SES students decreased, teacher use of autonomy-supportive strategies increased (0-10 %:  $M_4=20.3$ ,  $SD_4 = 3.2$ ; 11-25 %:  $M_4=20.7$ ,  $SD_4 = 3.1$ ; 26-50 %:  $M_4=20.7$ ,  $SD_4 = 3.1$ , More than 50 %:  $M_4=21.6$ ,  $SD_4=3.0$ ). The test yielded small effect sizes for all groups, as measured by the eta-squared, except for the one with moderate output-driven freedom from control ( $\eta^2 = 0.037$ , i.e., moderate).

Table 12.

*Classroom Practices by the Share of Disadvantaged Students*

	Country groups by types of teacher autonomy			
	Group 1: Limited	Group 2: Moderate Output-driven	Group 3: Moderate Input-driven	Group 4: Extended
Reading comprehension strategies	30.4 (4.6)	30.8 (4.6)	26.9 (5.0) <sup>1</sup>	26.7 (5.1)
0 -10 %	30.3 (4.9)	30.4 (5.3)	27.0 (5.1) <sup>1</sup>	27.5 (5.0)
11-25 %	31.1(4.2)	29.5 (5.4)	26.9 (5.3)	27.7 (4.7)
26-50%	29.7 (4.8)	28.1 (5.3)	27.2 (5.0)	29.5 (4.9)
More than 50 %				
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	3, 3187205 =	3, 1408688 =	3, 1687104 =	3, 342010 =
Eta-squared	9819.136	22000.274	227.850	3130.086
Eta-squared	0.009	0.045	0.000	0.027
Content differentiation				
0 -10 %	6.1 (1.5)	5.7 (1.6)	5.7 (1.3)	5.5 (1.4)
11-25 %	6.0 (1.5)	5.6 (1.6)	5.7 (1.5)	5.4 (1.4)
26-50%	5.9 (1.5)	5.3 (1.5)	5.8 (1.4)	5.3 (1.3)
More than 50 %	5.8 (1.6)	5.1 (1.7)	5.9 (1.5)	5.6 (1.4)
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	3, 3187205 =	3, 1408688 =	3, 1687104 =	3, 342010 =
Eta-squared	4435.217	11681.771	1397.404	360.265
Eta-squared	0.004	0.024	0.002	0.003
Autonomy-supportive strategies				
0 -10 %	21.8 (3.1)	23.8 (2.7)	21.0 (3.1)	20.3 (3.2)
11-25 %	21.5 (3.3)	24.1 (2.8)	20.9 (3.3)	20.6 (3.3)
26-50%	22.2 (3.2)	22.5 (3.5)	20.5 (3.3)	20.7 (3.1)
26-50%	21.3 (3.4)	22.9 (3.5)	20.6 (3.6)	21.6 (3.0)
More than 50 %				
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	3, 3187205 =	3, 1408688 =	3, 1687104 =	3, 342010 =
Eta-squared	97018.305	18155.275	1629.004	1523.067
Eta-squared	0.009	0.037	0.003	0.013

<sup>1</sup> - not significant (Difference between two groups (0-10 % And 11-25 %), *p*-value > 0.05))  
*p*-value (Two-way ANOVA) <0.001 for all analysis.

**Use of individualized instruction.** I used the chi-square test of association to examine the distribution of teachers by the frequency of using individualized instruction at schools with different share of low SES students across all four types of teacher

autonomy. It revealed a significant association between the use of individualized instruction and school composition in all groups of countries by autonomy ( $p < 0.001$ ). As Table 13 shows, in countries with extended teacher autonomy, as the share of students with low SES increased, the percentage of teachers using individualized instruction often or always also increased (0-10% - 33.0%, 11-25% - 36.1 %, 26-50% - 43.5%, more than 50% - 61.5%). Overall, the distribution was reversed in the group with moderate output-driven autonomy (0-10% - 51.0%, 11-25% - 54.2 %, 26-50% - 32.6%, more than 50% - 38.1%). Differences in the share of teachers were not large in countries with limited and moderate input-driven autonomy. In the latter, the share of teachers using individualized instruction was the highest at schools with 0-10 % and more than 50 % of low SES students (32.2 % and 38.7 %, respectively). In the systems with limited autonomy, for the most part, there were no differences in the share of teachers using this practice more or less frequently.

Table 13.

*Individualized Instruction by the Share of Disadvantaged Students*

	<b>Teachers Using Individualized Instruction (%)</b>		<i>p</i> -value
	<b>Never or Sometimes</b>	<b>Often or always</b>	
<b>Group 1: Limited</b>			
0-10%	55.5	44.5	<0.001
11-25%	48.6	51.4	
26-50%	40.0	60.0	
more than 50%	48.8	51.2	
<b>Group 2: Moderate output-driven</b>			
0-10%	49.0	51.0	<0.001
11-25%	45.8	54.2	
26-50%	67.4	32.6	
more than 50%	61.9	38.1	
<b>Group 3: Moderate input-driven</b>			
0-10%	67.8	32.2	<0.001
11-25%	77.8	22.2	

26-50%	73.1	26.9	
more than 50%	61.3	38.7	
Group 4: Extended			
0-10%	67.0	33.0	
11-25%	63.9	36.1	<0.001
26-50%	56.5	43.5	
more than 50%	38.5	61.5	

To summarize, in the countries with moderate output-driven autonomy, the higher the share of low SES students at schools, the less frequently teachers use all four practices in their classrooms. The pattern is similar, with smaller differences and few exceptions in the countries with limited teacher autonomy. The distribution of the frequency of teacher practices favored more disadvantaged schools best in countries with extended autonomy. As the share of students with low SES decreased at schools, teachers used these practices more frequently. The differences in the frequency of the use of effective classroom practices were the smallest in countries with moderate input-driven autonomy, yet, in most cases favoring schools with a higher share of disadvantaged students.

#### ***Teacher use of Classroom Practices by the Limitations Related to Student Attributes***

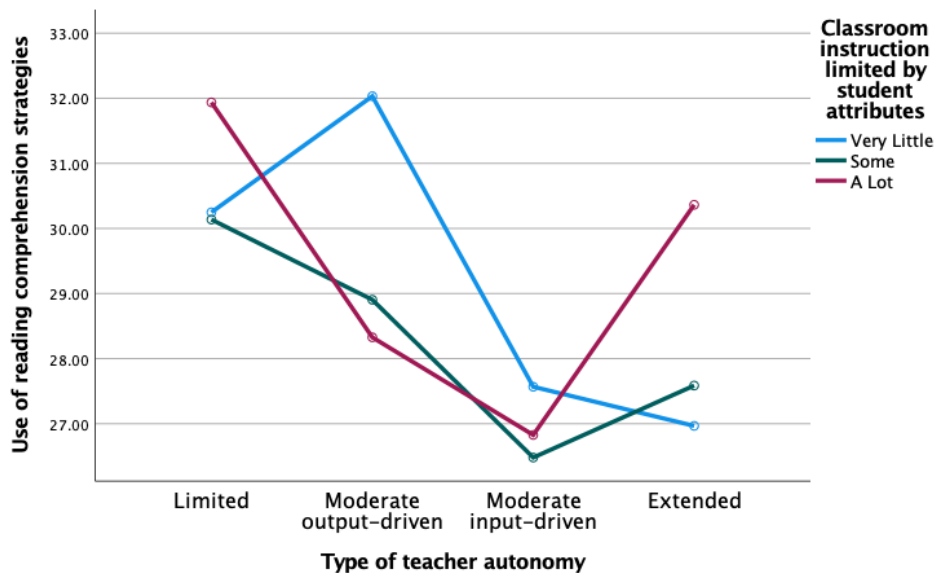
I also examined how teachers' application of selected classroom practices varied by teachers' perceptions of student-related attributes limiting their instruction. For this purpose, I used one-way ANOVA and two-way ANOVA, as well as a chi-square test (to analyze the use of individualized instruction).

**Use of reading comprehension strategies.** One-way ANOVA test revealed statistically significant differences ( $p < 0.001$ ) in the average use of strategies for reading comprehension between teachers with different perceptions about their instruction being limited by student attributes across all four groups of countries by teacher autonomy. The

multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of reading comprehension strategies between all categories of teachers within all four groups of countries by autonomy. As measured by eta-squared, the effect size was found to be small in groups of countries with limited ( $\eta^2 = 0.011$ ) and moderate input-driven ( $\eta^2 = 0.011$ ), and extended autonomy ( $\eta^2 = 0.008$ ). It was medium in a group with moderate output-driven autonomy ( $\eta^2 = 0.067$ ). In addition, a two-way ANOVA test detected a statistically significant interaction between teachers' perception of the limited instruction and type of autonomy in how they together associated with teacher use of reading comprehension strategies ( $F(6,6626475) = 17068.698, p < 0.001$ ).

Figure 9.

*Reading Comprehension Strategies by Teachers' Perceptions*



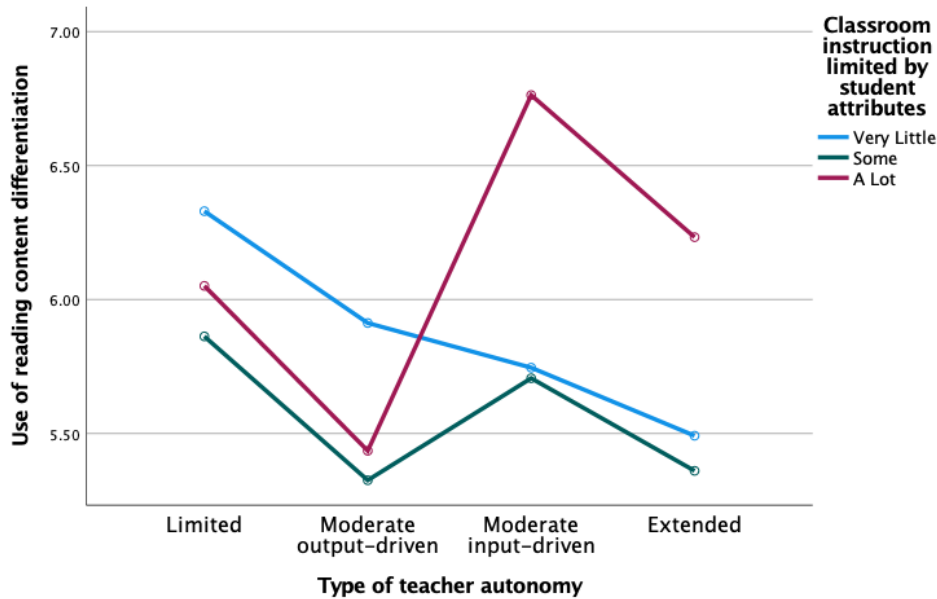
As depicted in Table 14 and Figure 9, in countries with limited and extended autonomy, teachers that perceive their instruction being significantly (i.e., a lot) limited by their student attributes used reading comprehension strategies more frequently

compared to other teachers (a lot:  $M_1 = 32.0$ ,  $SD_1 = 3.8$ ;  $M_4 = 30.3$ ,  $SD_4 = 3.4$ , Vs. Very little:  $M_1 = 30.3$ ,  $SD_1 = 4.8$ ;  $M_4 = 27.0$ ,  $SD_4 = 5.1$ ). This relationship was reversed in countries with moderate output-driven autonomy: teachers who thought that student attributes limit their instruction a lot used reading comprehension strategies least frequently ( $M_2 = 28.3$ ,  $SD_2 = 5.2$ ), compared to those perceiving this factor as a minor effect ( $M_2 = 32.0$ ,  $SD_2 = 4.2$ ). The relationship was similar, with a smaller magnitude of differences, in countries with moderate input-driven autonomy.

**Use of reading content differentiation.** One-way ANOVA test revealed statistically significant differences ( $p < 0.001$ ) in the average use of reading content differentiation between teachers with different perceptions about their instruction being limited by student attributes across all four groups of countries with different forms of autonomy. Also, the multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of these strategies between all categories of teachers with varying perceptions within all four groups of countries. A two-way ANOVA test detected a statistically significant interaction between teachers' perception of the limited instruction and type of freedom from control in how they together associated with teacher use of content differentiation ( $F(6, 6626475) = 8405.146$ ,  $p < 0.001$ )).

Figure 10.

*Reading Content Differentiation by Teachers' Perceptions*



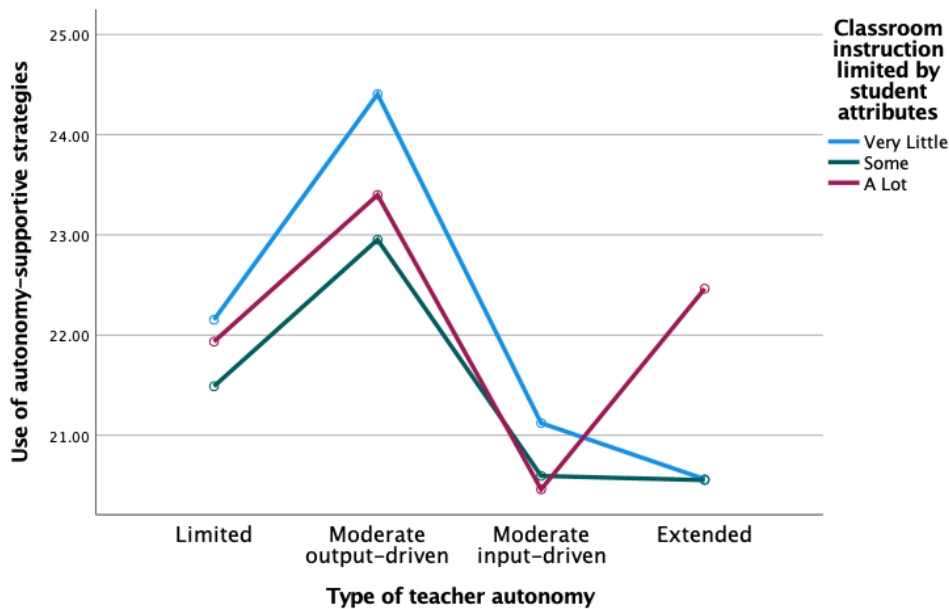
As depicted in Table 14 and Figure 10, in countries with moderate input-driven and extended autonomy, teachers that perceived their instruction being significantly (i.e., a lot) limited by their student attributes used reading content differentiation strategies more frequently compared to other teachers (a lot:  $M_3 = 6.7$ ,  $SD_3 = 1.4$ ;  $M_4 = 6.2$ ,  $SD_4 = 1.2$ , Very little:  $M_3 = 5.7$ ,  $SD_3 = 1.4$ ;  $M_4 = 5.5$ ,  $SD_4 = 1.4$ ). The pattern was reversed in countries with limited and moderate output-driven autonomy. On average, those teachers who thought their instruction was very little limited by student attributes, used content differentiation strategies most frequently (a lot:  $M_1 = 6.1$ ,  $SD_1 = 1.5$ ;  $M_2 = 5.4$ ,  $SD_2 = 1.6$ , Vs. Very little:  $M_1 = 6.3$ ,  $SD_1 = 1.5$ ;  $M_2 = 5.9$ ,  $SD_2 = 1.6$ ). However, the effect size measured by eta-squared was found to be small in all four groups of countries (limited:  $\eta^2 = 0.017$ ; moderate output-driven:  $\eta^2 = 0.023$ ; moderate input-driven:  $\eta^2 = 0.016$ , extended:  $\eta^2 = 0.006$ ).



**Use of autonomy-supportive strategies.** I also examined the distribution of the use of autonomy-supportive strategies given teachers' different perceptions about their instruction being limited by student attributes. The results of the one-way ANOVA test. The differences in the use of strategies by teachers with different perceptions of their instruction limited by student attributes, were found to be statistically significant ( $p < 0.001$ ). The multiple comparisons test (Bonferroni adjusted) revealed a statistically significant difference in the average use of autonomy-supportive strategies between all categories of teachers within all four groups of countries by teacher autonomy. A two-way ANOVA test revealed a statistically significant interaction between teachers' perceptions and type autonomy on teacher use of autonomy-supportive strategies ( $F(6, 6626475) = 3505.438, p < 0.001$ )).

Figure 11.

*Autonomy-supportive Strategies by Teachers' Perceptions*



Similar to the use of two other strategies, teachers who were granted extended autonomy and perceived their instruction as being limited by student attributes a lot, used autonomy-supportive strategies most frequently compared to the other two groups of teachers (A lot:  $M_4 = 22.5$ ,  $SD_4 = 3.0$ ; Some:  $M_4 = 20.5$ ,  $SD_4 = 3.2$ , A little:  $M_4 = 20.6$ ,  $SD_4 = 3.2$ ). In all other groups, teachers who thought that student attributes very little limited their instruction used autonomy-supportive strategies slightly, yet, most frequently (see Table 14 and Figure 11). It should be noted that effect size was found to be small in all groups (limited:  $\eta^2 = 0.004$ , moderate input-driven:  $\eta^2 = 0.006$ , extended:  $\eta^2 = 0.004$ ), except the one with moderate output-driven freedom from control ( $\eta^2 = 0.037$ ).

Table 14.

*Classroom Practices by Teachers' Perceptions*

	<b>Country groups by types of teacher autonomy</b>			
	<b>Group 1: Limited</b>	<b>Group 2: Moderate Output-driven</b>	<b>Group 3: Moderate Input-driven</b>	<b>Group 4: Extended</b>
Use of reading comprehension strategies	30.3 (4.8)	32.0 (4.2)	27.6 (5.3)	27.0 (5.1)
Very little	30.1 (4.7)	28.9 (5.3)	26.5 (4.8)	27.6 (5.1)
Some	32.0 (3.8)	28.3 (5.2)	26.8 (5.4)	30.3 (3.4)
A lot				
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	2, 3187206 = 17633.028	2, 1408689 = 50678.268	2, 1687105 = 9280.463	2, 342011 = 1310.472
Eta-squared	0.011	0.067	0.011	0.008
Use of content differentiation	6.3 (1.5)	5.9 (1.6)	5.7 (1.4)	5.5 (1.4)
Very little	5.9 (1.5)	5.4 (1.6)	5.7 (1.4)	5.4 (1.2)
Some	6.1 (1.5)	5.4 (1.6)	6.7 (1.4)	6.2 (1.2)
A lot				
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	2, 3187206 = 26799.028	2, 1408689 = 16781.744	2, 1687105 = 13383.024	2, 342011 = 1231.578
Eta-squared	0.017	0.023	0.016	0.006

Use of autonomy-supportive strategies	22.2 (3.2)	24.4 (2.6)	21.1 (3.3)	20.6 (3.2)
Very little	21.5 (3.3)	23.0 (3.4)	20.6 (3.3)	20.5 (3.2)
Some	21.9 (3.1)	23.4 (2.6)	20.5 (3.5)	22.5 (3.0)
A lot				
<i>p</i> -value	<0.001	<0.001	<0.001	<0.001
F Statistics	2, 3187206 =	2, 1408689 =	2, 1687105 =	2, 342011 =
	12700.805	26892.650	5491.448	740.309
Eta-squared	0.004	0.037	0.006	0.004
<i>p</i> -value (Two-way ANOVA) <0.001 for all analysis.				

**Use of individualized instruction.** I used the chi-square test of association to examine the distribution of teachers by the frequency of using individualized instruction across the groups with different types of autonomy by their perception of the instruction limited by student attributes (See Table 15). It revealed a significant association between the use of individualized instruction and teachers' perceptions in all groups ( $p < 0.001$ ). Overall, around 50 % of teachers with limited and moderate output-driven autonomy used individualized strategies often or always; however, in both groups of countries, this share decreased among the teachers who perceived their instruction to be limited by student attributes (40,6 % and 41.1 %, respectively). In countries with moderate input-driven autonomy, the share of teachers using individualized instruction did not vary by their perceptions. However, in this group, an overall smaller share of teachers (approximately 30 %) used individualized instruction. The distribution of teachers in countries with extended autonomy was different. Only 37.8 % of those who thought their instruction was very little limited used individualized instruction. This share got as high as 62.2% among teachers who believed that student attributes greatly limited their educational processes.

Table 15.

*Individualized Instruction by Teachers' Perceptions*

		<b>Teachers Using Individualized Instruction (%)</b>		<b>Pearson chi-square <i>p</i>-value</b>
		<b>Never or Sometimes</b>	<b>Often or Always</b>	
Group 1: Limited				
	Very little	48.4	52.6	
	Some	48.8	51.2	<0.001
	A lot	59.4	40.6	
Group 2: Moderate output-driven				
	Very little	43.8	56.2	
	Some	59.9	40.1	<0.001
	A lot	58.9	41.1	
Group 3: Moderate input-driven				
	Very little	67.9	32.1	
	Some	73.9	26.1	<0.001
	A lot	67.6	32.4	
Group 3: Extended				
	Very little	68.1	31.9	
	Some	56.8	43.2	<0.001
	A lot	37.8	62.2	

To summarize the findings on teachers' perceptions, in the countries with moderate output-driven autonomy, the frequency of using selected classroom practices decreased as the extent to which teachers thought student attributes limited their instruction increased. The pattern was similar, with a smaller magnitude of differences and exceptions with the use of reading comprehension strategies in countries with limited autonomy. This relationship was reversed in countries with extended autonomy: the more teachers thought student attributes limited their instruction, the more they used all four classroom practices. In countries with moderate input-driven autonomy, there was a smaller difference in the use of classroom practices, except for the use of reading content

differentiation. Yet, those teachers who thought that their instruction was very little limited by student attributes used these practices more frequently.

### ***Summary of the Findings***

Overall, findings suggest that teachers with limited and moderate output-driven autonomy used all four practices in their classrooms more frequently than those with moderate input-driven and extended ones. However, in the systems with moderate output-driven autonomy, teachers at schools with a higher composition of low SES students used these strategies less frequently than those engaged with more affluent students and schools. The pattern was similar, with a smaller magnitude of differences and few exceptions in the countries with limited autonomy. In systems with extended autonomy, teachers at schools with a higher share of low SES students used these classroom practices more frequently. Mostly, a similar pattern was observed in countries with moderate input-driven autonomy. Finally, only in countries with extended autonomy, teachers, who thought that student attributes limited their instruction, still used classroom practices more frequently than their peers, less concerned by the same limitations.

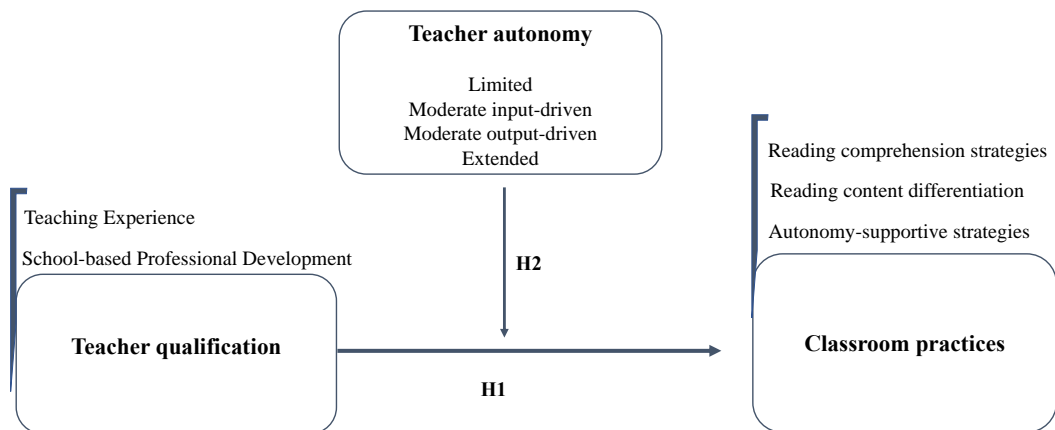
### **Teacher Autonomy as a Moderator: Teacher Qualification and Classroom Practices**

I used the multiple regression analysis with interaction variables to address the fourth research question, i.e., examine whether teacher autonomy moderated the relationship between teachers' qualification characteristics, such as teaching experience and school-based professional development, and the use of classroom practice (reading comprehension strategies, reading content differentiation, autonomy-supportive strategies). Specifically, as shown in Figure 12, I tested the following two hypotheses:

- Hypothesis 1: Teacher qualification characteristics predict the use of classroom practices so that a higher level of qualification is associated with the frequency of classroom practices on average.
- Hypothesis 2. The extent of teacher autonomy at a country level moderates the relationship between teacher qualifications and classroom practices.

Figure 12.

*An Analysis Model for Research Question 4*



Before testing my hypothesis, I used cross-tabulation to examine the distribution of teachers by each of the four qualification characteristics within each group by teacher autonomy. As Table 16 displays, in the group of countries with limited and extended autonomy, countries with extended autonomy have the lowest share of teachers with less than a BA education and the highest percentage of those with an MA or higher education. The share of the highest educational-level teachers is the lowest in countries with moderate output-driven autonomy. The distribution of teachers by the length of teaching experience is similar. Approximately half of the teachers have experience of 20 years or more. Also, Table 16 shows that teachers in countries with limited and output-driven

autonomy participate more intensively in formal professional development sessions.

These countries have a higher share of teachers participating in school-based professional activities often or very often.

Table 16.

*Teacher Qualification Characteristics by Teacher Autonomy*

<b>Teacher Qualification Characteristics</b>	<b>Group 1: Limited</b>	<b>Group 2: Moderate Output-driven</b>	<b>Group 3: Moderate Input-driven</b>	<b>Group 4: Extended</b>
%				
<b>Education</b>				
Less than BA	20.1	30.8	25.3	3.6
BA or equivalent	52.1	63.6	37.7	54.5
MA or higher	27.9	5.5	37.0	41.9
<b>Teaching experience</b>				
Less than 5 years	8.9	12.6	10.4	11.7
5 -9 years	13.2	13.3	14.1	13.2
10-19 years	23.5	25.7	33.5	28.0
20 years and more	54.4	48.4	42.0	47.1
<b>Formal professional development</b>				
None	17.0	13.7	28.3	29.9
Less than 6 hours	17.0	14.1	32.1	26.5
6-15 hours	21.8	16.0	22.6	25.3
16-35 hours	14.5	26.7	8.8	9.1
More than 35 hours	29.7	29.6	8.1	9.0
<b>School-based professional development</b>				
Never or almost never	4.2	2.1	6.1	3.3
Sometimes	35.0	26.0	57.4	49.1
Often	51.0	52.2	33.7	43.9
Very often	9.9	19.7	2.8	3.8

I conducted multiple linear regression analyses for each of the three classroom practices separately and used these practices as dependent variables (use of reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies). I used the following predictor variables:

- Teacher qualification characteristics (teaching experience and school-based professional development (a continuous version of the variable), one variable per analysis)).
- Teacher autonomy: (1) limited, (2) moderate output-driven, (3) moderate input-driven, (4) extended. I dummy-coded and used extended autonomy as a reference group.

In addition, I used the following control variables with indicated reference groups (when applicable):

- Classroom size (continuous).
- Instruction affected by resource shortages: (1) not affected (reference group), (2) somewhat affected, (3) affected a lot.
- Country's economic development level: (1) lower-middle income (reference group), (2) upper-middle income, (3) high-income.
- Country's democratic development level: (1) authoritarian regime (reference group), (2) hybrid regime, (3) flawed democracy, (4) full democracy.
- Teacher education: (1) less than BA (reference group), (2) BA or equivalent, (3) MA or higher.
- Teaching experience (continuous) (When applicable).
- Formal professional development: (1) none (reference group), (2) less than 6 hours, (3) 6-15 hours, (4) 16-35 hours, (5) more than 35 hours.
- School-based professional development: (1) never or almost never (reference group), (2) sometimes, (3) often, (4) very Often (When applicable).



For each of the above-mentioned analyses, I used four model specifications. I included control variables in the following order: all control variables (model 1), teacher qualification characteristics: teaching experience or school-based professional development) (model 2), teacher autonomy (model 3), and interaction variables (model 4).

Before describing the findings in response to the two hypotheses, I summarize the linear regression results in relation to the use of classroom practices across the groups by teacher autonomy. After controlling for all the variables mentioned above, these findings mostly confirmed the results of the descriptive analysis presented in response to research question 3. As shown by the results of model 3 (see Column 3 of Tables 17 through 22), teachers in the countries with limited autonomy used the reading comprehension strategies most frequently ( $\beta = 0.210, p = 0.000$ ). Educators in the countries with moderate output-driven ( $\beta = 0.021, p < 0.001$ ) and moderate input-driven autonomy ( $\beta = 0.014, p < 0.01$ ) also used them more frequently than those in the countries with extended autonomy (a reference group). Findings were quite identical for the use of reading content differentiation. Teachers with limited autonomy used them most frequently ( $\beta = 0.260, p = 0.000$ ). Educators in the countries with moderate output-driven ( $\beta = 0.115, p < 0.000$ ) and moderate input-driven autonomy ( $\beta = 0.124, p = 0.000$ ) also applied them more frequently than those with extended autonomy (a reference group). Findings of the use of autonomy-supportive strategies are also consistent with those of descriptive analysis: teachers in output-driven autonomy used these practices most frequently ( $\beta = 0.225, p = 0.000$ ), followed by those with limited ( $\beta = 0.092, p = 0.000$ ) and input-driven ( $\beta = 0.062, p = 0.000$ ). Thus, compared to the peers in all three groups, the reference

group, i.e., teachers in the extended autonomy, used the autonomy-supportive strategies less frequently.

### *Teacher Autonomy Moderating the Effect of Teaching Experience*

First, I estimated the predictive power of teaching experience in relation to **the use of reading comprehension strategies** (see Table 17). The results of the multiple regression yielded a statistically significant effect of teaching experience on the frequency of using reading comprehension strategies ( $\beta = 0.117, p < 0.00$ ). It also captured 0.12 % of the variation in the frequency of using this classroom practice ( $R^2$  change = 0.012, model 2). Thus, this finding supported my first hypothesis: teaching experience predicted the frequency of reading comprehension strategies, so more years of teaching experience was associated with more frequent use of reading comprehension strategies.

In the next step, I estimated the influence of the different types of autonomy on the relationship between teaching experience on the frequency of reading comprehension strategies (model 4). Compared to the extended autonomy, a reference group, moderate output-driven and input-driven autonomy, reinforced this effect ( $\beta = 0.085, p = 0.000$ , and  $\beta = 0.059, p = 0.000$ , correspondingly). Compared to all other groups, limited autonomy strengthened the relationship between experience and the selected classroom practice to the smallest extent ( $\beta = -0.039, p < 0.001$ ). To sum up, the evidence supported my hypothesis. Each form of teacher autonomy moderated the influence of teaching experience on reading comprehension strategies differently. The magnitude of the moderation was highest in the group of counties with moderate output-driven autonomy and moderate input-driven teacher autonomy and lowest in those with limited one.

Table 17.

*Autonomy Moderating the Effect of Teaching Experience on Reading Comprehension**Strategies*

Model Variables	Models in Regression Analysis			
	1	2	3	4
Adjusted R <sup>2</sup>	0.207	0.219	0.248	0.251
R <sup>2</sup> change	0.207	0.012	0.028	0.003
	Standardized Coefficients ( $\beta$ )			
Classroom size	0.013	0.015	0.012	0.015
Resource shortage (somewhat)	-0.064	-0.062	-0.043	-0.043
Resource shortage (a lot)	-0.024	-0.023	-0.007	-0.010
Economic development (upper-middle)	0.165	0.131	0.171	0.186
Democratic development (hybrid)	0.076	0.060	0.025	0.031
Democratic development (flawed)	0.111	0.086	0.173	0.172
Democratic development (full)	-0.080	-0.097	-0.040	-0.039
Teacher education (BA or eq.)	-0.019	0.013	-0.013	-0.004
Teacher education (MA or higher)	-0.006	0.016	-0.017	-0.006
Formal Professional Development (<6 hours)	0.059	0.057	0.048	0.047
Formal Professional Development (6-15 hours)	0.128	0.120	0.095	0.093
Formal Professional Development (16-35 hours)	0.145	0.142	0.128	0.126
Formal Professional Development (>35 hours)	0.245	0.230	0.194	0.192
School-based professional development	0.257	0.256	0.262	0.263
Teaching experience		0.117	0.097	0.076
Autonomy (limited)			0.210	0.235
Autonomy (mod. output-driven)			0.021	-0.060
Autonomy (mod. input-driven)			0.014	-0.033
Limited * Teaching experience				-0.039
Mod. output-driven * Teaching experience				0.085
Mod. input-driven * Teaching experience				0.059

Dependent variable: use of reading comprehension strategies.  
All coefficients are statistically significant at least with a  $p$ -value < 0.05 unless indicated otherwise.  
Economic development (high-income) was excluded from the regression analysis.

In the subsequent analysis, I estimated the relationship between teaching experience and the **use of reading content differentiation** (see Table 18). As results in Column 2 indicate, when controlling for all school, country-level, and teacher-level variables, including other qualification characteristics, an increase in the length of teaching experience was associated with more frequent use of reading content differentiation ( $\beta = 0.088$ ,  $p < 0.001$ ). It also accounted for 0.7 % of the variation in the frequency of using these strategies ( $R^2$  change = 0.007, model 2). Thus, the evidence

supported my hypothesis: the length of teaching experience predicted the use of content differentiation so that more years of experience increased the frequency of this classroom practice.

I also estimated how different types of teacher autonomy moderated the relationship between teaching experience and the use of content differentiation. As shown in Column 4 of Table 18, the effect of teaching experience was moderated differently by a statistically significant interaction between teaching experience and each type of teacher autonomy, capturing 0.4 % of the variation in the use of this classroom practice ( $R^2$  change = 0.004, model 4). Interaction coefficients were higher for all types of autonomy compared to the extended one. Among them, the highest were those of the interactions of teaching experience with limited ( $\beta = 0.166, p = 0.000$ ) and moderate output-driven ( $\beta = 0.174, p = 0.000$ ) autonomy. Compared to them, input-driven autonomy moderated the influence of the teaching experience on the use of content differentiation to a smaller extent ( $\beta = 0.052, p < 0.001$ ). Thus, the evidence supported my hypothesis. The effect of teaching experience on the use of content differentiation varied depending on the type of teacher autonomy. Specifically, experience had a more substantial positive effect on the use of content differentiation in countries with limited and moderate output-driven autonomy and the weakest in those with extended one.

Table 18.

*Autonomy Moderating the Effect of Teaching Experience on Reading Content*

*Differentiation*

Model Variables	Models in Regression Analysis			
	1	2	3	4
Adjusted R <sup>2</sup>	0.102	0.109	0.126	0.129
R <sup>2</sup> change	0.102	0.007	0.017	0.004

Model Variables	Models in Regression Analysis			
	1	2	3	4
	Standardized Coefficients ( $\beta$ )			
Classroom size	0.024	0.026	0.020	0.020
Resource shortage (somewhat)	-0.053	-0.052	-0.041	-0.040
Resource shortage (a lot)	-0.073	-0.072	-0.063	-0.065
Economic development (upper-middle)	-0.029	-0.055	-0.032	-0.039
Democratic development (hybrid)	0.078	0.066	0.043	0.040
Democratic development (flawed)	0.111	0.092	0.157	0.155
Democratic development (full)	0.179	0.166	0.205	0.198
Teacher education (BA or eq.)	0.019	0.043	0.033	0.032
Teacher education (MA or higher)	0.008	0.025	0.011	0.006
Formal Professional Development (<6 hours)	0.053	0.052	0.044	0.047
Formal Professional Development (6-15 hours)	0.098	0.092	0.074	0.080
Formal Professional Development (16-35 hours)	0.101	0.099	0.088	0.092
Formal Professional Development (>35 hours)	0.145	0.134	0.108	0.108
School-based professional development	0.245	0.245	0.251	0.253
Teaching experience		0.088	0.077	-0.051
Autonomy (limited)			0.260	0.138
Autonomy (mod. output-driven)			0.115	-0.038
Autonomy (mod. input-driven)			0.124	0.077
Limited * Teaching experience				0.166
Mod. output-driven * Teaching experience				0.174
Mod. input-driven * Teaching experience				0.052

Dependent variable: use of reading content differentiation.  
All coefficients are statistically significant at least with a  $p$ -value  $< 0.05$  unless indicated otherwise.  
Economic development (high-income) was excluded from the analysis.

In addition to the two classroom practices above, I also examined the moderating power of autonomy in relation to **the use of autonomy-supportive strategies** (see Table 19). In the first step, I estimated the relationship between teaching experience and these strategies. The results of the multiple regression showed a statistically significant effect of teaching experience on the frequency of the use of autonomy-supportive strategies ( $\beta = 0.033$ ,  $p = 0.000$ ,  $R^2$  change = 0.001, model 2). Thus, the findings supported my hypothesis: Longer teaching experience was associated with more frequent use of these strategies. It should be noted that compared to other classroom practices (reading comprehension and reading content differentiation), experience predicted the use of autonomy-supportive strategies with a smaller magnitude. Also, it explained the smaller percentage of variation in the use of this specific classroom practice.

In the next step, I estimated how different types of autonomy moderated the relationship between teaching experience and using autonomy-supportive strategies (model 4). Compared to the extended autonomy, a reference group, moderate output-driven and input-driven autonomy, reinforced this effect ( $\beta = 0.097, p = 0.000$ , and  $\beta = 0.030, p < 0.001$ , correspondingly). Compared to all other groups, limited autonomy strengthened the relationship between experience and the selected classroom practice to the smallest extent ( $\beta = -0.007, p < 0.001$ ). To sum up, the evidence supported my hypothesis. Each form of teacher autonomy moderated the influence of teaching experience on autonomy-supportive strategies differently. The magnitude of the moderation was highest in the group of counties with moderate output-driven autonomy, followed by the ones with moderate input-driven one. Limited and extended autonomy demonstrated the smallest moderating effect.

Table 19.

*Autonomy Moderating the Effect of Teaching Experience on Autonomy-supportive Strategies*

Model Variables	Models in Regression Analysis			
	1	2	3	4
Adjusted R <sup>2</sup>	0.193	0.194	0.209	0.211
R <sup>2</sup> change	0.193	0.001	0.015	0.002
	Standardized Coefficients ( $\beta$ )			
Classroom size	0.051	0.052	0.036	0.038
Resource shortage (somewhat)	-0.020	-0.019	-0.033	-0.032
Resource shortage (a lot)	0.012	0.012	-0.008	-0.011
Economic development (upper-middle)	0.048	0.038	-0.037	-0.028
Democratic development (hybrid)	-0.005	-0.010	0.002	0.005
Democratic development (flawed)	0.027	0.020	0.037	0.035
Democratic development (full)	-0.044	-0.048	-0.038	-0.040
Teacher education (BA or eq.)	-0.061	-0.052	-0.041	-0.035
Teacher education (MA or higher)	-0.083	-0.076	-0.032	-0.027
Formal Professional Development (<6 hours)	0.052	0.052	0.056	0.055
Formal Professional Development (6-15 hours)	0.096	0.093	0.098	0.099

Model Variables	Models in Regression Analysis			
	1	2	3	4
Formal Professional Development (16-35 hours)	0.167	0.166	0.156	0.155
Formal Professional Development (>35 hours)	0.203	0.199	0.206	0.205
School-based professional development	0.300	0.300	0.286	0.286
Teaching experience		0.033	0.057	0.029
Autonomy (limited)			0.092	0.094
Autonomy (mod. output-driven)			0.225	0.136
Autonomy (mod. input-driven)			0.062	0.037
Limited * Teaching experience				-0.007
Mod. output-driven * Teaching experience				0.097
Mod. input-driven * Teaching experience				0.030

Dependent variable: use of autonomy-supportive strategies.  
All coefficients are statistically significant at least with a  $p$ -value < 0.05 unless indicated otherwise.  
Economic development (high-income) was excluded from the analysis.

### ***Teacher Autonomy Moderating the Effect of School-based Professional Development***

In the subsequent analyses, I examined the moderating effect of teacher autonomy on the influence of school-based professional development, another teacher qualification characteristic. First, I estimated these relationships for **the use of reading comprehension strategies**. Multiple regression showed a statistically significant and high-magnitude effect of school-based professional development on the frequency of the use of reading comprehension strategies ( $\beta = 0.254$ ,  $p = 0.000$ , model 2). It also accounted for a 5.5 % variance in using these strategies ( $R^2$  change = 0.055). Thus, the findings supported my hypothesis: More frequent participation in school-based professional development activities was associated with more frequent use of these strategies (see Table 20).

However, as the interaction coefficients indicate in Column 4, this relationship was moderated differently by different types of teacher autonomy. Compared to that of the extended autonomy, interaction coefficients were lower and statistically significant for limited ( $\beta = -0.101$ ,  $p = 0.001$ ) and output-driven autonomy ( $\beta = -0.127$ ,  $p < 0.001$ )

and slightly higher for the input-driven one ( $\beta = 0.011, p = 0.001$ ). Also, as depicted in column 4, the inclusion of this two-way interaction accounted for 0.1 % variation in the use of reading comprehension strategies ( $R^2$  change = 0.001). To conclude, the findings supported my hypothesis. Each form of teacher autonomy moderated the influence of school-based professional development on reading comprehension strategies differently. Compared to extended and input-driven autonomy, limited and output-driven ones weakened the relationship between school-based professional development and the use of reading comprehension strategies. Among the first two, the moderate input-driven autonomy reinforced the examined relationship more strongly.

Table 20.

*Autonomy Moderating the Effect of School-based Professional Development on Reading Comprehension Strategies*

Model Variables	Models in Regression Analysis			
	1	2	3	4
Adjusted R <sup>2</sup>	0.165	0.219	0.248	0.251
R <sup>2</sup> change	0.165	0.055	0.028	0.001
	Standardized Coefficients ( $\beta$ )			
Classroom size	0.018	0.015	0.012	0.012
Resource shortage (somewhat)	-0.081	-0.062	-0.043	-0.042
Resource shortage (a lot)	-0.025	-0.023	-0.007	-0.007
Economic development (upper-middle)	0.101	0.131	0.171	0.170
Democratic development (hybrid)	0.013	0.060	0.025	0.022
Democratic development (flawed)	0.004	0.086	0.173	0.169
Democratic development (full)	-0.171	-0.097	-0.040	-0.043
Teacher education (BA or eq.)	-0.002	0.013	-0.013	-0.013
Teacher education (MA or higher)	0.001 <sup>1</sup>	0.016	-0.017	-0.016
Formal Professional Development (<6 hours)	0.067	0.057	0.048	0.048
Formal Professional Development (6-15 hours)	0.147	0.120	0.095	0.093
Formal Professional Development (16-35 hours)	0.185	0.142	0.128	0.128
Formal Professional Development (>35 hours)	0.294	0.230	0.194	0.195
Teaching experience	0.118	0.117	0.097	0.096
School-based professional development		0.256	0.262	0.299
Autonomy (limited)			0.210	0.300
Autonomy (mod. output-driven)			0.021	0.132
Autonomy (mod. input-driven)			0.014	0.007 <sup>1</sup>
Limited * School-based professional development				-0.101



Model Variables	Models in Regression Analysis			
	1	2	3	4
Mod. output-driven * School-based professional development				-0.127
Mod. input-driven * School-based professional development				0.011

Dependent variable: use of reading comprehension strategies.  
All coefficients are statistically significant at least with a  $p$ -value < 0.05 unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
Economic development (high-income) was excluded from the analysis.

I next estimated the relationship between school-based professional development and **the use of reading content differentiation** (see Table 21). As results in Column 2 indicate, when controlling for all school, country-level, and teacher-level variables, including other qualification characteristics, an increase in the intensity of the school-based professional development activities was associated with more frequent use of reading content differentiation ( $\beta = 0.245, p = 0.000$ ). It also accounted for 5.0 of the variation in the frequency of using this classroom practice ( $R^2$  change = 0.050). Thus, the findings supported the first hypothesis: school-based professional development predicted the use of content differentiation so that more intensity of school-based professional development increased the frequency of this classroom practice.

I also estimated how different types of teacher autonomy moderated the relationship between school-based professional development and the use of content differentiation (model 4). As shown in Column 4 of Table 21, the effect of school-based professional development was moderated differently by a statistically significant interaction between these school-based activities and each type of teacher autonomy, capturing 0.4 % of the variation in the use of this classroom practice ( $R^2$  change = 0.004, model 4). Interaction coefficients were higher for all types of autonomy compared to the extended one. Among them, the highest were those of the interactions of school-based

professional development with moderate output-driven ( $\beta = 0.430, p = 0.000$ ) and limited ( $\beta = 0.180, p = 0.000$ ) autonomy. Compared to them, input-driven autonomy moderated the influence of the teaching experience on the use of content differentiation to a smaller extent ( $\beta = 0.076, p < 0.001$ ). Thus, the evidence supported my hypothesis. The effect of school-based professional development on the use of content differentiation varied depending on the type of teacher autonomy. The magnitude of the moderation was highest in the group of counties with moderate output-driven autonomy and limited teacher autonomy and lowest in those with input-driven and extended ones.

Table 21.

*Autonomy Moderating the Effect of School-based Professional Development on Reading Content Differentiation*

Model Variables	Models in Regression Analysis			
	1	2	3	4
Adjusted R <sup>2</sup>	0.059	0.109	0.126	0.130
R <sup>2</sup> change	0.059	0.050	0.017	0.004
	Standardized Coefficients ( $\beta$ )			
Classroom size	0.028	0.026	0.020	0.019
Resource shortage (somewhat)	-0.070	-0.052	-0.041	-0.042
Resource shortage (a lot)	-0.074	-0.072	-0.063	-0.064
Economic development (upper-middle)	-0.083	-0.055	-0.032	-0.036
Democratic development (hybrid)	0.020	0.066	0.043	0.040
Democratic development (flawed)	0.014	0.092	0.157	0.157
Democratic development (full)	0.095	0.166	0.205	0.203
Teacher education (BA or eq.)	0.028	0.043	0.033	0.037
Teacher education (MA or higher)	0.010	0.025	0.011	0.015
Formal Professional Development (<6 hours)	0.061	0.052	0.044	0.045
Formal Professional Development (6-15 hours)	0.118	0.092	0.074	0.076
Formal Professional Development (16-35 hours)	0.140	0.099	0.088	0.087
Formal Professional Development (>35 hours)	0.195	0.134	0.108	0.106
Teaching experience	0.089	0.088	0.077	0.080
School-based professional development		0.245	0.251	0.148
Autonomy (limited)			0.260	0.106
Autonomy (mod. output-driven)			0.115	-0.269
Autonomy (mod. input-driven)			0.124	0.040
Limited * School-based professional development				0.180
Mod. output-driven * School-based professional development				0.430
Mod. input-driven * School-based professional development				0.076

Model Variables	Models in Regression Analysis			
	1	2	3	4
Dependent variable: use of reading content differentiation.				
All coefficients are statistically significant at least with a $p$ -value $< 0.05$ unless indicated otherwise.				
Economic development (high-income) was excluded from the analysis.				

Findings on **using autonomy-supportive strategies** were quite identical to those on reading comprehension strategies (see Table 22). Multiple regression showed a statistically significant and high-magnitude effect of school-based professional development on the frequency of using autonomy-supportive strategies ( $\beta = 0.300$ ,  $p = 0.000$ , model 2). It also captured a high 7.5 % variance in using these strategies ( $R^2$  change = 0.075). Thus, the findings supported my hypothesis: More frequent participation in school-based professional development activities was associated with more frequent use of autonomy-supportive strategies.

I also estimated how different types of teacher autonomy moderated the relationship between school-based professional development and the use of content differentiation (model 4). Compared to that of the extended autonomy, interaction coefficients were lower and statistically significant for limited ( $\beta = -0.105$ ,  $p < 0.001$ ) and output-driven autonomy ( $\beta = -0.143$ ,  $p = 0.000$ ). Compared to the same reference variable, it was slightly higher for the input-driven one ( $\beta = 0.019$ ,  $p < 0.001$ ). Also, as depicted in column 4, the inclusion of this two-way interaction accounted for 0.1 % variation in the use of reading comprehension strategies ( $R^2$  change = 0.001). To conclude, the findings supported my hypothesis. Each form of teacher autonomy moderated the influence of school-based professional development autonomy-supportive strategies differently. Compared to extended and input-driven autonomy, limited and output-driven ones weakened the relationship between school-based professional

development and the use of autonomy-supportive strategies. Among the first two, the moderate input-driven autonomy reinforced the examined relationship more strongly.

Table 22.

*Autonomy Moderating the Effect of School-based Professional Development on  
Autonomy-supportive Strategies*

Model Variables	Models in Regression Analysis				
		1	2	3	4
Adjusted R <sup>2</sup>	0.119	0.194	0.209	0.210	
R <sup>2</sup> change	0.119	0.075	0.015	0.001	
	Standardized Coefficients (β)				
Classroom size	0.056	.052	0.036	0.036	
Resource shortage (somewhat)	-0.042	-.019	-0.033	-0.032	
Resource shortage (a lot)	0.010	.012	-0.008	-0.008	
Economic development (upper-middle)	0.003	.038	-0.037	-0.038	
Democratic development (hybrid)	-0.065	-.010	0.002	-0.001 <sup>1</sup>	
Democratic development (flawed)	-0.076	.020	0.037	0.033	
Democratic development (full)	-0.135	-.048	-0.038	-0.041	
Teacher education (BA or eq.)	-0.070	-0.052	-0.041	-0.040	
Teacher education (MA or higher)	-0.095	-0.076	-0.032	-0.032	
Formal Professional Development (<6 hours)	0.063	0.052	0.056	0.055	
Formal Professional Development (6-15 hours)	0.125	0.093	0.098	0.096	
Formal Professional Development (16-35 hours)	0.217	0.166	0.156	0.156	
Formal Professional Development (>35 hours)	0.273	0.199	0.206	0.207	
Teaching experience	0.034	0.033	0.057	0.056	
School-based professional development		0.300	0.286	0.324	
Autonomy (limited)			0.092	0.185	
Autonomy (mod. output-driven)			0.225	0.352	
Autonomy (mod. input-driven)			0.062	0.047	
Limited * School-based professional development				-0.105	
Mod. output-driven * School-based professional development				-0.143	
Mod. input-driven * School-based professional development				0.019	

Dependent variable: use of autonomy-supportive strategies.  
 All coefficients are statistically significant at least with a *p*-value < 0.05 unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
 Economic development (high-income) was excluded from the analysis.

**Summary of the Findings**

As the findings above across two teacher qualification characteristics and three teacher classroom practices demonstrate, teaching experience and school-based

professional development are positively and significantly associated with reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies. This indicates that the more years of teaching experience teachers had and the more intensively they participated in school-based professional development, the more frequently they used all three classroom practices with their students. Compared to teaching experience, participation in school-based professional development was the strongest predictor magnitude-wise for all three classroom practices. Among these practices, it was most strongly associated with using autonomy-supportive strategies and least with reading content differentiation. Unlike this, teaching experience demonstrated the weakest predictive power in relation to autonomy-supportive strategies and the strongest – reading comprehension strategies.

In addition, I found that the relationship between classroom practices and each qualification characteristic varied depending on the type of teacher autonomy at the national level. Compared to other forms of autonomy, moderate output-driven one the most strongly reinforced the predictive power teaching experience. Stated differently, with the same years of teaching experience, teachers that were granted output-driven autonomy from the national government used all three practices most frequently. Moderate input-driven autonomy also quite strongly strengthened the relationship between teaching experience and these strategies. Compared to others, limited and extended autonomy had the slightest effect. Among the two, limited autonomy had the weakest moderating power in relation to reading comprehension and autonomy-supportive strategies, while for reading content differentiation, teaching years mattered least in countries with extended autonomy.

The moderating pattern of teacher autonomy was different for school-based professional development. The findings suggest that limited and output-driven autonomy strongly reinforced the relationship between school-based professional development and reading content differentiation. At the same time, they had the smallest moderating effect in relation to reading comprehension and autonomy-supportive strategies. Compared to others, input-driven autonomy consistently strengthened the effectiveness of school-based professional development for all three classroom practices.

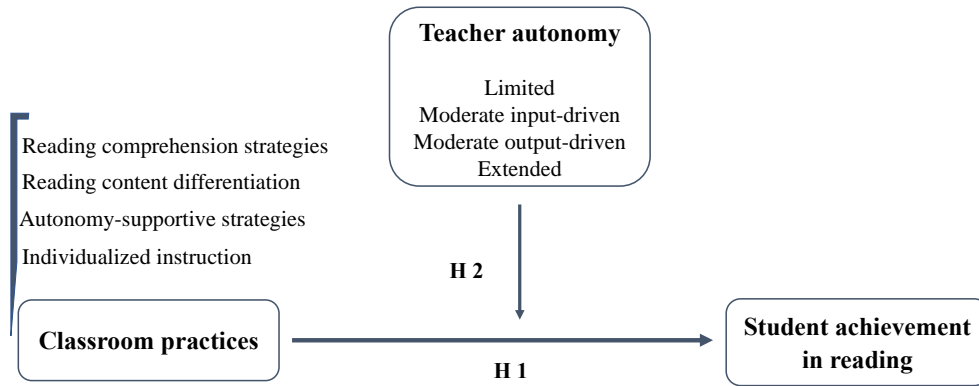
### **Teacher Autonomy as a Moderator: Classroom Practices and Student Achievement**

I used the multiple regression analysis with interaction variables to examine whether teacher autonomy moderated the relationship between teachers' use of classroom practice and student achievement scores in reading. Specifically, as shown in Figure 13, I tested the following two hypotheses:

- Hypothesis 1: Classroom practice predicts student achievement in reading so that a higher frequency of usage is associated with higher student scores on average.
- Hypothesis 2: The extent of teacher autonomy at a country level moderates the relationship between classroom practices and student achievement.

Figure 13.

*An Analysis Model for Research Question 5*



Due to the high pair-wise correlations between the three practices (see Table 23), I conducted separate analyses for each of the four classroom practices.

Table 23.

*Intercorrelation Results for Classroom Practices*

Classroom Practices	1	2	3
1 Use of reading comprehension strategies		0.369	0.401
2 Use of reading content differentiation	0.369		0.638
3 Use of autonomy-supportive strategies	0.401	0.638	

All coefficients are statistically significant at least with a  $p$ -value  $<0.05$  unless indicated otherwise.

N=6,625,025

As shown in Tables 24 through 27, I used student reading score as the dependent variable in all regression tests. I used the following predictor variables:

- Classroom practice (reading comprehension strategies (continuous), content differentiation (continuous), autonomy-supportive strategies (continuous), and individualized instruction (0 – never or sometimes, 1 – often or always)).

- Teacher autonomy: (1) limited, (2) moderate output-driven, (3) moderate input-driven, (4) extended. I dummy-coded and used extended autonomy as a reference group.

In addition, I used the following control variables with indicated reference groups (when applicable):

- Student gender: (0) boy, (1) girl.
- Student socioeconomic background (continuous).
- Classroom size (continuous).
- Instruction affected by resource shortages: (1) not affected (reference group), (2) somewhat affected, (3) affected a lot.
- Country's economic development level: (1) lower-middle-income (reference group), (2) upper-middle-income, (3) high-income-country).
- Country's democratic development level: (1) authoritarian regime (reference group), (2) hybrid regime, (3) flawed democracy, (4) full democracy.
- Teacher education: (1) less than BA (reference group), (2) BA or equivalent, (3) MA or higher.
- Teaching experience (continuous).
- Formal professional development: (1) none (reference group), (2) less than 6 hours, (3) 6-15 hours, (4) 16-35 hours, (5) more than 35 hours.
- School-based professional development: (1) never or almost never (reference group), (2) sometimes, (3) often, (4) very Often.

I conducted multiple regression analyses for each classroom practice with seven model specifications. I included control variables in the following order: students'



socioeconomic background and gender (model 1), classroom size and shortage of reading resources (model 2), country's economic and democratic development (model 3), and teacher qualification characteristics (model 4). In the remaining models, I included the variables of primary interest in the following order: the use of reading comprehension strategies (model 5), teacher autonomy (model 6), and interaction variables (model 7).

Below I summarize the findings of models 5 through 7 for each classroom practice as a primary predictor variable. As I will explain, based on the results obtained from models 5 and 7, I accepted or rejected my hypotheses 1 and 2, respectively.

### ***Use of Reading Comprehension Strategies and Student Achievement in Reading***

To test the first hypothesis, I estimated the predictive power of reading comprehension strategies in relation to student achievement (See Table 24). As shown in Column 5, after controlling for the above-mentioned student, teacher, school, and country-level variables, an increase in the frequency of the use of these strategies was associated with an increase in student reading score ( $\beta = 0.117$ ;  $p < 0.001$ ,  $R^2 = 0.011$ ). In other words, I found support for my first hypothesis in the data.

I also examined how different types of autonomy moderated the relationship between reading comprehension strategies and student reading scores (model 7). To estimate the moderation effect, I included an interaction variable between the frequency of use of reading comprehension strategies and each of the teacher autonomy-type indicators. All three types of autonomy reinforced the relationship between the use of reading comprehension strategies and student achievement in reading compared to the extended autonomy, which served as a reference group. The coefficient of the interaction with reading comprehension strategies was highest for the moderate output-driven

autonomy ( $\beta = 0.308, p < 0.001$ ), followed by the moderate input-driven autonomy ( $\beta = 0.157, p < 0.001$ ), and limited autonomy ( $\beta = 0.095, p < 0.001$ ). To sum up, the evidence supported my hypothesis. Each form of teacher autonomy moderated the influence of the classroom practice on student achievement differently. The magnitude of the moderation was highest in the group of counties with moderate output-driven autonomy and moderate input-driven teacher autonomy and lowest in those with extended and limited ones.

Table 24.

*Autonomy as a Moderator: Reading Comprehension Strategies and Student Achievement*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.342	0.361	0.411	0.431	0.442	0.486	0.487
R <sup>2</sup> change	0.342	0.019	0.050	0.020	0.011	0.044	0.001
	Standardized Coefficients ( $\beta$ )						
Student SES	0.575	0.551	0.476	0.445	0.439	0.388	0.385
Student gender (girl)	0.091	0.092	0.097	0.095	0.093	0.091	0.092
Classroom size		-0.107	-0.061	-0.047	-0.049	-0.021	-0.018
Resource shortage (somewhat)		-0.076	-0.066	-0.060	-0.054	-0.037	-0.036
Resource shortage (a lot)		-0.083	-0.071	-0.066	-0.064	-0.034	-0.035
Economic development (upper-middle)			0.296	0.246	0.231	0.367	0.377
Democratic development (hybrid)			-0.065	-0.080	-0.088	-0.126	-0.129
Democratic development (flawed)			0.238	0.212	0.205	0.203	0.206
Democratic development (full)			0.217	0.196	0.211	0.211	0.210
Teacher education (BA or eq.)				0.026	0.025	0.013	0.014
Teacher education (MA or higher)				0.094	0.094	0.031	0.033
Teaching experience				0.096	0.083	0.050	0.048
Formal Professional Development (<6 hours)				0.013	0.006	0.001 <sup>1</sup>	-0.000 <sup>1</sup>
Formal Professional Development (6-15 hours)				0.023	0.008	0.001 <sup>1</sup>	-0.000 <sup>1</sup>
Formal Professional Development (16-35 hours)				-0.021	-0.038	-0.019	-0.019
Formal Professional Development (>35 hours)				0.041	0.013	0.004	0.003
School-based professional development (sometimes)				0.024	0.012	0.010	0.013
School-based professional development (often)				0.056	0.021	0.028	0.031
School-based professional development (very Often)				-0.016	-0.043	-0.012	-0.010
Comprehension strategies					0.117	0.093	0.028

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Autonomy (limited)						-0.038	-0.117
Autonomy (mod. output-driven)						-0.307	-0.606
Autonomy (mod. input-driven)						-0.019	-0.175
Comprehension strategies * Limited							0.095
Comprehension strategies * Mod. output-driven							0.308
Comprehension strategies * Mod. input-driven							0.157

Dependent variable: student score in reading.  
All coefficients are statistically significant at least with a  $p$ -value  $< 0.05$  unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
Economic development (high-income) was excluded from the analysis.

### *Use of Reading Content Differentiation and Student Achievement in Reading*

After examining the moderating power of autonomy in relation to reading comprehension strategies, I conducted a similar analysis for another classroom practice – reading content differentiation (Table 25). As shown in Column 5, when controlling for the above-mentioned student, teacher, school, and country-level variables as part of the first four models, an increase in the frequency of the use of these strategies was associated with a higher student reading score ( $\beta = 0.048$ ;  $p < 0.001$ ,  $R^2 = 0.002$ ). In other words, I found support for my first hypothesis.

In the next step, I estimated how different types of teacher autonomy at a country level moderated the relationship between content differentiation and student achievement in reading (model 7). Comparison with the extended autonomy, a reference group, shows that limited and moderate output-driven autonomy reinforced the relationship between teachers' reading content differentiation and student achievement. Among the two, the interaction coefficient was higher for the moderated output-driven autonomy ( $\beta = 0.149$ ,  $p < 0.001$ ) compared to the limited one ( $\beta = 0.063$ ,  $p < 0.001$ ). The results showed no statistically significant difference in how moderate input-driven autonomy, compared to

extended autonomy, affected the relationship between teacher use of content differentiation and student achievement in reading ( $\beta = 0.002$ ,  $p > 0.05$ ). To conclude, the findings supported my hypothesis partially. Input-driven and extended autonomy had a similar influence, while limited and output-driven ones more strongly reinforced the effect of the selected classroom practice.

Table 25.

*Autonomy as a Moderator: Reading Content Differentiation and Student Achievement*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.342	0.361	0.411	0.431	0.433	0.480	0.482
R <sup>2</sup> change	0.342	0.019	0.050	0.020	0.002	0.047	0.001
	Standardized Coefficients ( $\beta$ )						
Student SES	0.575	0.551	0.476	0.445	0.441	0.389	0.386
Student gender (girl)	0.091	0.092	0.097	0.095	0.096	0.093	0.094
Classroom size		-0.107	-0.061	-0.047	-0.048	-0.020	-0.019
Resource shortage (somewhat)		-0.076	-0.066	-0.060	-0.058	-0.040	-0.039
Resource shortage (a lot)		-0.083	-0.071	-0.066	-0.063	-0.032	-0.032
Economic development (upper-middle)			0.296	0.246	0.248	0.382	0.386
Democratic development (hybrid)			-0.065	-0.080	-0.084	-0.126	-0.126
Democratic development (flawed)			0.238	0.212	0.208	0.212	0.213
Democratic development (full)			0.217	0.196	0.189	0.299	0.202
Teacher education (BA or eq.)				0.026	0.024	0.010	0.012
Teacher education (MA or higher)				0.094	0.094	0.029	0.029
Teaching experience				0.096	0.092	0.056	0.054
Formal Professional Development (<6 hours)				0.013	0.010	0.004	0.004
Formal Professional Development (6-15 hours)				0.023	0.008	0.008	0.008
Formal Professional Development (16-35 hours)				-0.021	-0.026	-0.010	-0.009
Formal Professional Development (>35 hours)				0.041	0.035	0.019	0.018
School-based professional development (sometimes)				0.024	0.021	0.018	0.018
School-based professional development (often)				0.056	0.045	0.048	0.047
School-based professional development (very Often)				-0.016	-0.028	0.001	-0.002
Content differentiation					0.048	0.034	-0.004
Autonomy (limited)						-0.027	-0.081
Autonomy (mod. output-driven)						-0.307	-0.451
Autonomy (mod. input-driven)						-0.023	-0.002
Content differentiation * Limited							0.063

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Content differentiation * Mod. output-driven							0.149
Content differentiation * Mod. input-driven							0.002 <sup>1</sup>

Dependent variable: student score in reading.  
All coefficients are statistically significant at least with a  $p$ -value  $< 0.05$  unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
Economic development (high-income) was excluded from the analysis.

### *Use of Autonomy-supportive Strategies and Student Achievement in Reading*

In the subsequent analysis, in the first step, I estimated the predictive power of autonomy-supportive strategies in relation to student achievement (Table 26). As Column 5 indicates, when controlling for the above-mentioned student, teacher, school, and country-level variables as part of the first four models, an increase in the frequency of the use of these strategies was associated with higher student reading score, on average ( $\beta = 0.012$ ;  $p < 0.001$ ,  $R^2 = 0.001$ ). This finding supported my first hypothesis.

In the next step, I estimated the influence of different types of teacher autonomy on the relationship between the use of autonomy-supportive strategies and student achievement in reading (model 7). As Table 26 depicts, all three types of autonomy reinforced the relationship between the use of reading comprehension strategies and student achievement in reading compared to the extended autonomy, a reference group. The coefficient of the interaction with autonomy-supportive strategies was highest for the moderate output-driven autonomy ( $\beta = 0.199$ ,  $p < 0.001$ ), followed by the moderate input-driven autonomy ( $\beta = 0.151$ ,  $p < 0.001$ ) and limited autonomy ( $\beta = 0.052$ ,  $p < 0.001$ ). To sum up, the evidence supported my hypothesis. Each form of teacher autonomy moderated the influence of the classroom practice on student achievement differently. The magnitude of the moderation was highest in the group of counties with

moderate output-driven autonomy and moderate input-driven teacher autonomy and lowest in those with extended and limited ones.

Table 26.

*Autonomy as a Moderator: Autonomy-supportive Strategies and Student Achievement*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.342	.361	0.411	0.431	0.431	0.481	0.482
R <sup>2</sup> change	0.342	.019	0.050	0.020	0.000	0.050	0.001
	Standardized Coefficients ( $\beta$ )						
Student SES	0.575	0.551	0.476	0.445	0.445	0.390	0.389
Student gender (girl)	0.091	0.092	0.097	0.095	0.095	0.092	0.092
Classroom size		-0.107	-0.061	-0.047	-0.048	-0.021	-0.021
Resource shortage (somewhat)		-0.076	-0.066	-0.060	-0.060	-0.039	-0.039
Resource shortage (a lot)		-0.083	-0.071	-0.066	-0.066	-0.034	-0.034
Economic development (upper-middle)			0.296	0.246	0.245	0.383	0.382
Democratic development (hybrid)			-0.065	-0.080	-0.080	-0.124	-0.126
Democratic development (flawed)			0.238	0.212	0.212	0.217	0.216
Democratic development (full)			0.217	0.196	0.197	0.208	0.206
Teacher education (BA or eq.)				0.026	0.026	0.013	0.015
Teacher education (MA or higher)				0.094	0.095	0.031	0.033
Teaching experience				0.096	0.095	0.056	0.055
Formal Professional Development (<6 hours)				0.013	0.012	0.003	0.002
Formal Professional Development (6-15 hours)				0.023	0.022	0.005	0.004
Formal Professional Development (16-35 hours)				-0.021	-0.023	-0.014	-0.015
Formal Professional Development (>35 hours)				0.041	0.039	0.012	0.011
School-based professional development (sometimes)				0.024	0.023	0.014	0.014
School-based professional development (often)				0.056	0.052	0.040	0.041
School-based professional development (very Often)				-0.016	-0.019	0.003	-0.003
Autonomy-supportive strategies					0.012	0.048	0.012
Autonomy (limited)						-0.023	-0.067
Autonomy (mod. output-driven)						-0.315	-0.499
Autonomy (mod. input-driven)						-0.021	-0.168
Autonomy-supportive strategies * Limited							0.052
Autonomy-supportive strategies * Mod. output-driven							0.199
Autonomy-supportive strategies * Mod. input-driven							0.151

Dependent variable: student score in reading.

All coefficients are statistically significant at least with a *p*-value < 0.05 unless indicated otherwise.

<sup>1</sup> - indicates non-significant coefficients.

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Economic development (high-income) was excluded from the analysis.							

### *Use of Individualized Instruction and Student Achievement in Reading*

To understand the relationship between individualized instruction and student achievement, I included the use of individualized instruction as a dummy variable (0 – never or sometimes, 1 – often or always). As results in Table 27 indicate, when controlling for the above-mentioned student, teacher, school, and country-level variables, students whose teachers used individualized instruction often or always scored lower in reading than those whose teachers less frequently (never or sometimes) used the same approaches in their classrooms ( $\beta = -0.014, p < 0.001$ ).

I also estimated how different types of teacher autonomy moderated the relationship between the use of individualized instruction and student achievement in reading (model 7). As shown in Table 27, compared to the extended autonomy (a reference group), limited and input-driven autonomy weakened the negative relationship between the use of individualized instruction and student achievement ( $\beta = -0.003, p < 0.001, \beta = -0.023, p < 0.001$ , respectively). However, as reflected in the coefficient, the difference in how extended and moderate input-driven autonomy influenced the above-mentioned relationship was statistically significant but of minimal magnitude. Compared to all other types of autonomy, moderate output-driven autonomy strengthened the negative relationship between individualized instruction and student achievement. To conclude, the findings supported my hypothesis. Each type of autonomy moderated the relationship between individualized instruction and student reading scores differently. In other words, compared to other forms of autonomy, individualized instruction had a

weaker negative impact on student achievement in countries with limited autonomy, while this impact was stronger in those with moderate output-driven autonomy.

Table 27.

*Teacher Autonomy as a Moderator: Individualized Instruction and Student Achievement*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.342	0.361	0.411	0.431	0.431	0.479	0.480
R <sup>2</sup> change	0.342	0.019	0.050	0.020	0.000	0.048	0.001
	Standardized Coefficients ( $\beta$ )						
Student SES	0.575	0.551	0.476	0.445	0.444	0.391	0.390
Student gender (girl)	0.091	0.092	0.097	0.095	0.095	0.093	0.093
Classroom size		-0.107	-0.061	-0.047	-0.047	-0.020	-0.019
Resource shortage (somewhat)		-0.076	-0.066	-0.060	-0.060	-0.041	-0.040
Resource shortage (a lot)		-0.083	-0.071	-0.066	-0.066	-0.035	-0.035
Economic development (upper-middle)			0.296	0.246	0.248	0.381	0.381
Democratic development (hybrid)			-0.065	-0.080	-0.081	-0.124	-0.126
Democratic development (flawed)			0.238	0.212	0.214	0.217	0.216
Democratic development (full)			0.217	0.196	0.199	0.206	0.204
Teacher education (BA or eq.)				0.026	0.024	0.011	0.010
Teacher education (MA or higher)				0.094	0.093	0.029	0.028
Teaching experience				0.096	0.095	0.058	0.059
Formal Professional Development (<6 hours)				0.013	0.012	0.006	0.005
Formal Professional Development (6-15 hours)				0.023	0.022	0.010	0.009
Formal Professional Development (16-35 hours)				-0.021	-0.022	-0.007	-0.006
Formal Professional Development (>35 hours)				0.041	0.039	0.022	0.023
School-based professional development (sometimes)				0.024	0.024	0.020	0.019
School-based professional development (often)				0.056	0.054	0.055	0.054
School-based professional development (very Often)				-0.016	-0.018	0.010	0.009
Individualized Instruction (Often/always)					-0.014	0.001	0.008
Autonomy (limited)						-0.018	-0.007
Autonomy (mod. output-driven)						-0.303	-0.319
Autonomy (mod. input-driven)						-0.019	-0.017
Individualized Instruction * Limited							-0.023
Individualized Instruction *Mod. output-driven							0.022
Individualized Instruction*Mod. input-driven							-0.003

Dependent variable: student score in reading.

All coefficients are statistically significant at least with a *p*-value < 0.05 unless indicated otherwise.

<sup>1</sup> - indicates non-significant coefficients.



Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Economic development (high-income) was excluded from the analysis.							

### *Summary of the Findings*

As the findings demonstrate, all classroom practices except for individualized instruction (reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies) are positively and significantly associated with student achievement. Stated differently, the more frequently teachers use these strategies, the higher their students scored in reading assessments. Among the strategies, reading comprehension was the strongest predictor magnitude-wise, and autonomy strategies had the weakest. I also found that teacher autonomy at a country level had a different effect on the relationship between each classroom practice and student achievement. Compared to other forms of autonomy, a moderate output-driven one showed the strongest reinforcing power for the effectiveness of all strategies. Moderate input-driven autonomy also strongly strengthened the relationship between the use of reading comprehension and autonomy-supportive strategies. Compared to others, limited and extended autonomy had the slightest effect. Among the two, using the three classroom practices was less effective in the counties with extended autonomy. In other words, with the same frequency of the three strategies, students in countries with extended autonomy scored the lowest.

Unlike other classroom practices, results indicated a negative relationship between the frequent use of individualized instruction and student achievement in reading. Compared to other forms of autonomy, individualized instruction had the weakest negative impact on student achievement in countries with limited autonomy, while this impact was strongest in those with moderate output-driven autonomy. In other

words, with the same frequency of individualized instruction, students in countries with limited autonomy scored the highest, and those with moderate output-driven teacher autonomy - the lowest.

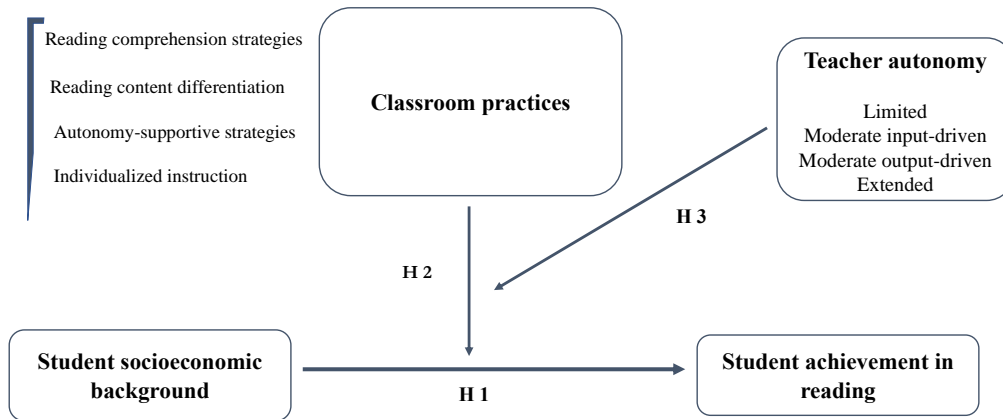
### **Teacher Autonomy as a Moderator for Student Equity**

I conducted a moderated moderation analysis to examine how teacher autonomy affects the influence of teacher practices on the relationship between student socioeconomic background and student achievement in reading. Specifically, as shown in Figure 14, I tested the following three hypotheses:

- Hypothesis 1. Student socioeconomic background strongly predicts student achievement in reading, i.e., students with higher values of socioeconomic background variable in the dataset are predicted to have higher reading scores.
- Hypothesis 2. Classroom practices moderate the relationship between student socioeconomic background and student achievement in reading so that socioeconomic background becomes a weaker predictor of student achievement.
- Hypothesis 3. The extent of teacher autonomy at a country level moderates the relationship between student socioeconomic background and student achievement.

Figure 14.

*An Analysis Model for Research Question 6*



I conducted a multiple regression analysis with two-way and three-way interactions to examine the primary and secondary moderating roles of classroom practices and teacher autonomy. As shown in Tables 28 through 31, I used student reading score (continuous measure) as the dependent variable and student socioeconomic background (continuous measure) as a predictor variable in all regression tests. I used each of the classroom practices as a primary moderator variable (reading comprehension strategies (continuous), content differentiation (continuous), autonomy-supportive strategies (continuous), and individualized instruction (0 – never or sometimes, 1 – often or always)). Teacher autonomy at a country level (coded as (1) limited, (2) moderate output-driven, (3) moderate input-driven, (4) extended)) was included as a secondary moderator variable. In addition, I used the following control variables with indicated reference groups for categorical variables:

- Student gender: (0) boy (reference), (1) girl.

- Classroom size (continuous)
- Instruction affected by resource shortages: (1) not affected (reference group), (2) somewhat affected, (3) affected a lot.
- Country's economic development level: (1) lower-middle-income (reference group), (2) upper-middle-income, (3) high-income-country).
- Country's democratic development level: (1) authoritarian regime (reference group), (2) hybrid regime, (3) flawed democracy, (4) full democracy.
- Teacher education: (1) less than BA (reference group), (2) BA or equivalent, (3) MA or higher.
- Teaching experience (continuous).
- Formal professional development: (1) none (reference group), (2) less than 6 hours, (3) 6-15 hours, (4) 16-35 hours, (5) more than 35 hours.
- School-based professional development: (1) never or almost never (reference group), (2) sometimes, (3) often, (4) very Often.

As shown in Tables 28-31, I did not include all four classroom practices in the regression simultaneously because these variables are highly correlated, which would result in multicollinearity. For each classroom practice, as a primary moderator variable, I conducted a separate multiple regression analysis sequentially including seven blocks of variables: these blocks included control variables (model 1), predictor variable (model 2), primary moderator variable (model 3), two-way interaction variable for predictor and moderator (Model 4), secondary moderator variable – dummy-coded (model 5), two-way interaction variables with the secondary moderator variable (model 6), three-way interactions (model 7).

Interpretation of the findings of model 2, which is common for all regression analyses for each of the classroom practices (see Tables 28-31), allowed me to test the hypothesized relationship between student socio-economic background and achievement in reading. After controlling for all student, school, and country-level variables, student socio-economic background still showed a significant positive impact on student reading scores ( $\beta = 0.445, p < 0.001$ ). Also, the change in R-square almost doubled the preceding one (from 0.289 to 0.431). In other words, student socioeconomic background explained nearly as much variation in achievement as the control variables included in Model 1 altogether. Thus, I found support for hypothesis 1: Students with one standard deviation increase in the value of socio-economic background score 0.445 standardized points higher, i.e., student socioeconomic background strongly predicted student achievement in reading.

Below I present the findings for Models 4 and 7 for each selected four classroom practices.

### ***Teacher Autonomy Moderating the Effect of Reading Comprehension Strategies***

To test hypothesis 2 about the moderating effect of the frequency of using reading comprehension strategies, I examined standardized coefficients in model 4 (See Table 28). Although I anticipated that the use of reading comprehension strategies would diminish the effect of student socio-economic background on student achievement, the findings showed that the two-way interaction (SES \* Reading comprehension strategies) was positive ( $\beta = 0.018, p < 0.001$ ). The coefficient of the socioeconomic background was also higher ( $\beta = 0.452, p < 0.001$ ) compared to those in previous models ( $\beta = 0.445, p < 0.001$  in model 2;  $\beta = 0.439, p < 0.001$  in model 3). Also, including the moderator did

not explain any variation in student score ( $R^2$  change = 0.000, model 4). To sum up, I did not find support for hypothesis 2. After adding the interaction effect of using reading comprehension strategies, the predictive power of the socio-economic background increased. In other words, the higher the socioeconomic status was and the more the teachers used reading comprehension strategies, the higher students scored. Thus, using reading comprehension strategies widened the achievement gap in reading between high and low SES students.

I used model 7 to address the third hypothesis, i.e., to examine how teacher autonomy - the secondary moderator – reinforced the moderating effect of reading comprehension strategies on the relationship between student socioeconomic background and achievement in reading. As the results indicate, student socioeconomic background remained a strong predictor of student achievement in reading ( $\beta = 0.357, p < 0.001$ , model 7). However, the coefficients of the three-way interaction (SES \* Reading comprehension strategies \* autonomy) varied by the type of autonomy. It was negative for extended autonomy, a reference group ( $\beta = -0.022, p < 0.001$ ). This means that extended autonomy mitigated the joint effect of socioeconomic status and reading comprehension strategies. For other types of autonomy, the three way-interaction effects were positive:  $\beta = 0.055, p < 0.001$  (limited),  $\beta = 0.057, p < 0.001$  (moderate output-driven), and  $\beta = 0.051, p < 0.001$  (moderate input-driven). In other words, compared to extended autonomy, they reinforced the effect of the joint effect of socioeconomic status and reading comprehension strategies, i.e., with more frequent use of reading comprehension strategies the students with higher SES scored even higher in reading. It should be noted that the addition of this three-way moderation explained no variation in

student achievement ( $R^2$  change = 0.000, model 7). In conclusion, I found support for hypothesis 3: Overall, teacher autonomy minimally reinforced the moderating effect of reading comprehension strategies on the relationship between student socioeconomic background and student achievement in reading. Among the four types of teacher autonomy, use of reading comprehension strategies showed to benefit lower SES students more in the countries with extended autonomy. This means that extended autonomy showed the potential to reduce the gap in reading achievement.

Table 28.

*Autonomy Moderating the Effect of Reading Comprehension Strategies*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.289	0.431	0.442	0.442	0.486	0.490	0.490
R <sup>2</sup> change	0.289	0.142	0.011	0.000	0.044	0.004	0.000
	Standardized coefficients ( $\beta$ )						
Student gender (girl)	0.104	0.095	0.093	0.093	0.091	0.090	0.090
Classroom size	-0.013	-0.047	-0.049	-0.049	-0.021	-0.017	-0.017
Resource shortages (somewhat affected)	-0.111	-0.060	-0.054	-0.053	-0.037	-0.035	-0.035
Resource shortages (affected a lot)	-0.109	-0.066	-0.064	-0.064	-0.034	-0.031	-0.031
Economic development (upper middle-income)	0.229	0.246	0.231	0.231	0.367	0.388	0.388
Democratic development (hybrid)	-0.218	-0.080	-0.088	-0.087	-0.126	-0.135	-0.135
Democratic development (flawed democracy)	0.274	0.212	0.205	0.205	0.203	0.206	0.206
Democratic development (full democracy)	0.298	0.196	0.211	0.210	0.211	0.220	0.220
Teacher education (BA or equivalent)	0.076	0.026	0.025	0.025	0.013	0.015	0.015
Teacher education (MA or higher)	0.182	0.094	0.094	0.094	0.031	0.033	0.033
Teaching experience	0.143	0.096	0.083	0.083	0.050	0.044	0.044
Formal professional development (<6 hours)	0.015	0.013	0.006	0.006	0.001	-0.001 <sup>1</sup>	-0.001
Formal professional development (6-15 hours)	0.024	0.023	0.008	0.008	0.001	0.000 <sup>1</sup>	0.000 <sup>1</sup>
Formal professional development (16-35 hours)	-0.031	-0.021	-0.038	-0.038	-0.019	-0.018 <sup>1</sup>	-0.018
Formal professional development (>35 hours)	0.052	0.041	0.013	0.014	0.004	0.003	0.003

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
School-based Professional Development (sometimes)	0.036	0.024	0.012	0.012	0.010	0.016	0.016
School-based Professional Development (often)	0.079	0.046	0.021	0.020	0.028	0.033	0.033
School-based Professional Development (very Often)	-0.021	-0.016	-0.043	-0.044	-0.012	-0.009	-0.009
Student SES		0.445	0.439	0.452	0.397	0.381	0.357
Comprehension strategies			0.117	0.107	0.086	0.026	0.042
Comprehension strategies * Student SES				0.018	0.013	0.010	-0.022
Autonomy (limited)					-0.038	-0.077	-0.143
Autonomy (mod. output-driven)					-0.307	-0.782	-0.845
Autonomy (mod. input-driven)					-0.020	-0.028	-0.089
Student SES * limited						-0.037	0.024
Student SES * mod. output-driven						0.226	0.279
Student SES * mod. input-driven						-0.132	-0.073
Comprehension strategies * limited						0.086	0.041
Comprehension strategies * mod. output-driven						0.252	0.207
Comprehension strategies * mod. input-driven						0.139	0.102
Comprehension strategies* Student SES * limited							0.055
Comprehension strategies * Student SES * mod. output-driven							0.057
Comprehension strategies * Student SES * mod. input-driven							0.041

Dependent variable: student score in reading.  
Standardized coefficients (betas) reported.  
All coefficients are statistically significant at least with a p-value < 0.05 unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
Economic development (high-income) was excluded from the regression analysis.

### ***Teacher Autonomy Moderating the Effect of Reading Content Differentiation***

Similarly, I examined standardized coefficients in model 4 (see Table 29) to estimate the moderating effect of the frequency of using reading content differentiation. According to the findings, use of reading content differentiation reinforced the effect of SES on student achievement ( $\beta = 0.033$ ,  $p < 0.001$ ,  $R^2$  change = 0.000). The standardized



coefficient of the socioeconomic background variable in model 4 also got higher ( $\beta = 0.461, p = 0.000$ ) than those in previous Models ( $\beta = 0.445, p = 0.000$  in model 2;  $\beta = 0.441, p < 0.001$  in model 3). To conclude, after adding the interaction effect of reading content differentiation, the predictive power of the socio-economic background increased. In other words, similar to reading comprehension strategies, using reading content differentiation increased the achievement gap in reading between high and low SES students. Thus, the findings did not support Hypothesis 2.

To test hypothesis 3, I examined the secondary moderating influence of teacher autonomy through model 7. With the inclusion of the three-way interactions, student socioeconomic background remained a strong predictor of student achievement in reading ( $\beta = 0.367, p < 0.001$ , model 7). The results indicated varying effects of the three-way interaction (SES \* Reading content differentiation \* autonomy) by type of teacher autonomy. The three-way-interaction coefficient was negative for extended autonomy, a reference group ( $\beta = -0.011, p < 0.001$ ). This means that it mitigated the joint of effect of socioeconomic status and reading comprehension strategies. The three-way interaction coefficients were not significant for moderate output-driven and moderate input-driven autonomy ( $\beta = 0.008, p > 0.05$  and  $\beta = 0.007, p > 0.05$ , respectively). This means these two types of autonomy had the same effect as the extended one. Finally, the coefficient for the limited autonomy was positive ( $\beta = 0.029, p < 0.001$ ). In other words, compared to extended autonomy, they reinforced the effect of the joint of effect of socioeconomic status and reading content differentiation, i.e., with more frequent use of reading content differentiation strategies the students with higher

SES scored even higher in reading. Finally, the inclusion of the three-way moderation did not account for any additional variation ( $R^2$  change = 0.000).

Thus, overall, teacher autonomy had small and not economically significant moderating effect in addition to teacher practices on the relationship between student socioeconomic background and student achievement in reading. All type of teacher autonomy, except for limited one, diminished the combined impact of student socioeconomic background and use of reading comprehension strategies on student achievement. This means that they all showed the potential to reduce the achievement gap. Limited autonomy positively reinforced the effect, i.e., increased the achievement gap. Thus, the findings satisfied hypothesis 3.

Table 29.

*Autonomy Moderating the Effect of Reading Content Differentiation*

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Adjusted R <sup>2</sup>	0.289	0.431	0.433	0.433	0.480	0.485	0.485
R <sup>2</sup> change	0.289	0.142	0.002	0.000	0.047	0.005	0.000
	Standardized coefficients ( $\beta$ )						
Student gender (girl)	0.104	0.095	0.096	0.096	0.093	0.092	0.092
Classroom size	-0.013	-0.047	-0.048	-0.048	-0.020	-0.017	-0.017
Resource shortages (somewhat affected)	-0.111	-0.060	-0.058	-0.058	-0.040	-0.038	-0.038
Resource shortages (affected a lot)	-0.109	-0.066	-0.063	-0.062	-0.032	-0.028	-0.028
Economic development (upper middle-income)	0.229	0.246	0.248	0.248	0.382	0.399	0.398
Democratic development (hybrid)	-0.218	-0.080	-0.084	-0.084	-0.126	-0.134	-0.134
Democratic development (flawed democracy)	0.274	0.212	0.208	0.208	0.212	0.213	0.213
Democratic development (full democracy)	0.298	0.196	0.189	0.189	0.199	0.212	0.212
Teacher education (BA or equivalent)	0.076	0.026	0.024	0.025	0.011	0.012	0.012
Teacher education (MA or higher)	0.182	0.094	0.094	0.093	0.029	0.029	0.029
Teaching experience	0.143	0.096	0.092	0.092	0.056	0.049	0.049
Formal professional development (<6 hours)	0.015	0.013	0.010	0.010	0.004	0.003	0.003

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Formal professional development (6-15 hours)	0.024	0.023	0.018	0.018	0.008	0.007	0.007
Formal professional development (16-35 hours)	-0.031	-0.021	-0.026	-0.026	-0.010	-0.008	-0.008
Formal professional development (>35 hours)	0.052	0.041	0.035	0.034	0.018	0.017	0.017
School-based Professional Development (sometimes)	0.036	0.024	0.021	0.021	0.018	0.022	0.022
School-based Professional Development (often)	0.079	0.056	0.045	0.045	0.048	0.049	0.049
School-based Professional Development (very Often)	-0.021	-0.016	-0.028	-0.028	0.001	0.000 <sup>1</sup>	0.000 <sup>1</sup>
Student SES		0.445	0.441	0.461	0.399	0.374	0.367
Content Differentiation			0.048	0.023	0.021	-0.005	0.003 <sup>1</sup>
Content Differentiation * Student SES				0.033	0.017	0.001 <sup>1</sup>	-0.011
Autonomy (limited)					-0.027	-0.051	-0.082
Autonomy (mod. output-driven)					-0.307	-0.684	-0.686
Autonomy (mod. input-driven)					-0.023	0.121	0.112
Student SES * limited						-0.042	-0.012 <sup>1</sup>
Student SES * mod. output-driven						0.250	0.259
Student SES * mod. input-driven						-0.150	-0.140
Content Differentiation. * Limited						0.068	0.043
Content Differentiation * Mod. output-driven						0.125	0.120
Content Differentiation * Mod. input-driven						0.005	-0.001 <sup>1</sup>
Content Differentiation * Student SES * limited							0.029
Content Differentiation * Student SES * mod. output-driven							0.008 <sup>1</sup>
Content Differentiation * Student SES * mod. input-driven							0.007 <sup>1</sup>

Dependent variable: student score in reading.  
Standardized coefficients (betas) reported.  
All coefficients are statistically significant at least with a *p*-value < 0.05 unless indicated otherwise.  
<sup>1</sup> - indicates non-significant coefficients.  
Economic development (high-income) was excluded from the analysis.

### ***Teacher Autonomy Moderating the Effect of Autonomy-supportive Strategies***

The results of the multiple regression analysis for the autonomy-supportive strategies as a primary moderator are presented in Table 30. Model 4 shows that similar to two other classroom practices, using autonomy-supportive strategies reinforces the relationship between student socio-economic background and student achievement ( $\beta =$

0.013,  $p < 0.001$ ). Correspondingly, the standardized coefficient of the socioeconomic background was also higher ( $\beta = 0.454$ ,  $p < 0.001$ ) compared to those in previous models ( $\beta = 0.445$ ,  $p < 0.001$  in model 2;  $\beta = 0.445$ ,  $p = 0.000$  in model 3). Yet, this interaction between socioeconomic status and the use of autonomy-supportive strategies did not account for any variation in student achievement ( $R^2$  change = 0.000). To sum up, I did not find support for hypothesis 2. Although I hypothesized that teachers' use of autonomy-supportive strategies would diminish the effect of socioeconomic background on student achievement, in fact, it increased its predictive power. In other words, more extensive use of autonomy-supportive strategies widened the achievement gap in reading between high and low-SES students.

I analyzed standardized coefficients from model 7 to respond to hypothesis 3, i.e., examine how teacher autonomy - the secondary moderator - influences the moderating effect of autonomy-supportive strategies on the relationship between student socioeconomic background and achievement in reading. As reflected in the coefficient, student socioeconomic background remained a strong predictor of student achievement in reading ( $\beta = 0.362$ ,  $p < 0.001$ , model 7). However, the effect of the three-way interaction (SES \* autonomy-supportive strategies \* autonomy) varied by the type of autonomy. As indicated in the results, it was negative for extended autonomy only ( $\beta = -0.015$ ,  $p < 0.001$ ). This means the use of autonomy-supportive strategies was more effective for lower SES students, i.e., extended autonomy mitigated the joint effect of socioeconomic status and reading comprehension strategies. For other types of autonomy, the three way-interaction effects were positive, i.e., compared to extended autonomy, they reinforced the moderated influence of autonomy-supportive strategies and socioeconomic

background on student achievement:  $\beta = 0.049, p < 0.001$  (limited),  $\beta = 0.094, p < 0.01$  (moderate output-driven), and  $\beta = 0.020, p < 0.001$  (moderate input-driven) ( $R^2$  change= 0.000).

To sum up, the results supported hypothesis 3. Each form of teacher autonomy at a country level impacted the moderating effect of autonomy-supportive strategies on the relationship between student socioeconomic background and achievement differently. Extended autonomy mitigated the joint influence of student socioeconomic background and the use of autonomy-supportive strategies on student achievement. Contrary to this, in the countries with limited, moderate input-driven and moderate output-driven autonomy, students with higher socioeconomic background and more frequent exposure to autonomy-supportive strategies scored even higher. The implication of this finding is that compared to extended autonomy, all three types of autonomy contributed to the increased gap in student achievement (with the lower magnitude in the case of moderate input-driven autonomy).

Table 30.

*Autonomy Moderating the Effect of Autonomy-supportive Strategies*

Model Variables	Models in the Regression Analysis						
		1	2	3	4	5	6
Adjusted R <sup>2</sup>	0.289	0.431	0.431	0.431	0.481	0.486	0.486
R <sup>2</sup> change	0.240	0.171	0.000	0.000	0.065	0.005	0.000
	Standardized coefficients ( $\beta$ )						
Student gender (girl)	0.104	0.095	0.095	0.095	0.092	0.091	0.091
Classroom size	-0.013	-0.047	-0.048	-0.048	-0.022	-0.019	-0.019
Resource shortages (somewhat affected)	-0.111	-0.060	-0.060	-0.060	-0.039	-0.037	-0.037
Resource shortages (affected a lot)	-0.109	-0.066	-0.066	-0.068	-0.034	-0.030	-0.030
Economic development (upper middle-income)	0.229	0.246	0.245	0.245	0.383	0.396	0.396
Democratic development (hybrid)	-0.218	0.080	-0.080	-0.080	-0.124	-0.133	-0.133
Democratic development (flawed democracy)	0.274	0.212	0.212	0.212	0.217	0.216	0.216

Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Democratic development (full democracy)	0.298	0.196	0.197	0.197	0.208	0.217	0.217
Teacher education (BA or equivalent)	0.076	0.026	0.026	0.026	0.014	0.015	0.016
Teacher education (MA or higher)	0.182	0.094	0.095	0.095	0.031	0.033	0.033
Teaching experience	0.143	0.096	0.096	0.095	0.056	0.050	0.050
Formal professional development (<6 hours)	0.015	0.013	0.012	0.012	0.003	0.001	0.001
Formal professional development (6-15 hours)	0.024	0.023	0.022	0.022	0.005	0.003	0.003
Formal professional development (16-35 hours)	-0.031	-0.021	-0.023	-0.023	-0.014	-0.014	-0.014
Formal professional development (>35 hours)	0.052	0.041	0.039	0.039	0.012	0.011	0.010
School-based Professional Development (sometimes)	0.036	0.024	0.023	0.022	0.014	0.018	0.018
School-based Professional Development (often)	0.079	0.056	0.052	0.052	0.040	0.042	0.042
School-based Professional Development (very Often)	-0.021	-0.016	-0.019	-0.019	-0.003	-0.002	-0.002
Student SES		0.445	0.445	0.454	0.397	0.386	0.362
Autonomy-supportive strategies			0.012	0.005	0.043	0.007	0.021
Autonomy-supportive strategies * Student SES				0.013	0.011	0.016	-0.015
Autonomy (limited)					-0.023	-0.029	-0.091
Autonomy (mod. output-driven)					-0.315	-0.711	-0.803
Autonomy (mod. input-driven)					-0.021	-0.014	-0.045
Student SES * limited						-0.041	0.016
Student SES * mod. output-driven						0.260	0.344
Student SES * mod. input-driven						-0.146	-0.116
Autonomy-supportive strategies * limited						0.046	0.007 <sup>1</sup>
Autonomy-supportive strategies * mod. output-driven						.0142	0.060
Autonomy-supportive strategies * mod. input-driven						-0.138	0.121
Autonomy-supportive strategies * Student SES * limited							0.049
Autonomy-supportive strategies* Student SES * mod. output-driven							0.094
Autonomy-supportive strategies* Student SES * mod. input-driven							0.020

Dependent variable: student score in reading.

All coefficients are statistically significant at least with a  $p$ -value < 0.05 unless indicated otherwise.

<sup>1</sup> - indicates non-significant coefficients.

Economic development (high-income) was excluded from the regression analysis.

### ***Teacher Autonomy Moderating the Effect of Individualized Instruction***

I examined standardized coefficients in model 4 (see Column 4 of Table 31) to estimate the moderating effect of the frequency of using reading content differentiation. According to the findings, the two-way interaction (SES \* individualized instruction) was negative ( $\beta = -0.030, p < 0.001, R^2 \text{ change} = 0.000$ ). The standardized coefficient of the socioeconomic background in model 4 also decreased ( $\beta = 0.439, p < 0.001$ ) compared to previous models ( $\beta = 0.445, p < 0.001$  in model 2 and  $\beta = 0.444, p < 0.001$  in model 3). To conclude, unlike the other three classroom practices, individualized instruction moderated the relationship between student socioeconomic background and student achievement in reading so that socioeconomic background became a weaker predictor of student achievement. Thus, the findings supported hypothesis 2.

To test hypothesis 3, I examined the secondary moderating influence of teacher autonomy through model 7 of the analysis (see Column 7 of Table 31). As indicated by the results, the effects of the three-way interaction (SES \* Reading content differentiation \* autonomy) varied by type of teacher autonomy. As reflected in the standardized coefficients, the three way-interaction was negative for extended autonomy, a reference group ( $\beta = -0.008, p < 0.001$ ). This means that individualized instruction was more effective for lower SES students. The coefficient of the input-driven autonomy was also negative ( $\beta = -0.018, p < 0.001$ ). This indicates that compared to extended autonomy, in the countries with moderate input-driven autonomy, the use of individualized instruction was even more effective for lower SES students. In contrast, positive coefficients of limited ( $\beta = 0.043, p < 0.001$ ) and moderate output-driven ( $\beta = 0.118, p = 0.001$ ) autonomy indicated that compared to the extended autonomy, the use of these strategies

benefited more students with higher SES. This means that they contributed to the increased gap in reading achievement.

Table 31.

*Autonomy Moderating the Effect of Individualized Instruction*

Model Variables	Models in the Regression Analysis						
		1	2	3	4	5	6
Adjusted R <sup>2</sup>	0.289	0.431	0.431	0.431	0.479	0.484	0.485
R <sup>2</sup> change	0.289	0.142	0.000	0.000	0.048	0.005	0.000
	Standardized coefficients ( $\beta$ )						
Student gender (girl)	0.104	0.095	0.095	0.096	0.093	0.092	0.092
Classroom size	-0.013	-0.047	-0.047	-0.047	-0.020	-0.017	-0.017
Resource shortages (somewhat affected)	-0.111	-0.060	-0.060	-0.060	-0.041	-0.038	-0.038
Resource shortages (affected a lot)	-0.109	-0.066	-0.066	-0.066	-0.035	-0.031	-0.031
Economic development (upper middle-income)	0.229	0.246	0.248	0.247	0.380	0.394	0.394
Democratic development (hybrid)	-0.218	-0.080	-0.081	-0.081	-0.124	-0.135	-0.135
Democratic development (flawed democracy)	0.274	0.212	0.214	0.214	0.217	0.216	0.216
Democratic development (full democracy)	0.298	0.196	0.199	0.198	0.205	0.215	0.216
Teacher education (BA or equivalent)	0.076	0.026	0.024	0.024	0.011	0.011	0.011
Teacher education (MA or higher)	0.182	0.094	0.093	0.093	0.029	0.028	0.029
Teaching experience	0.143	0.096	0.095	0.095	0.058	0.053	0.053
Formal professional development (<6 hours)	0.015	0.013	0.012	0.012	0.005	0.004	0.003
Formal professional development (6-15 hours)	0.024	0.023	0.022	0.022	0.010	0.008	0.007
Formal professional development (16-35 hours)	-0.031	-0.021	-0.022	-0.022	-0.007	-0.006	-0.006
Formal professional development (>35 hours)	0.052	0.041	0.039	0.039	0.022	0.021	0.020
School-based professional development (sometimes)	0.036	0.024	0.024	0.024	0.020	0.023	0.023
School-based professional development (often)	0.079	0.056	0.054	0.054	0.055	0.055	0.055
School-based professional development (very Often)	-0.021	-0.016	-0.018	-0.018	0.010	0.009	0.010
Student SES		0.445	0.444	0.439	0.385	0.365	0.373
Individualized instruction			0.014	0.042	0.032	0.046	0.016
Individualized instruction * Student SES				-0.030	-0.032	-0.044	-0.008
Autonomy (limited)					-0.018	.021	0.048
Autonomy (mod. output-driven)					-0.303	-.598	-0.529
Autonomy (mod. input-driven)					-0.019	0.121	0.120
Student SES * limited						-0.038	-0.064



Model Variables	Models in the Regression Analysis						
	1	2	3	4	5	6	7
Student SES * mod. output-driven						0.271	0.202
Student SES * mod. input-driven						-0.143	-0.141
Individualized instruction * limited						-0.019	-0.065
Individualized instruction * mod. output-driven						0.024	-0.096
Individualized instruction * mod. input-driven						-0.003	0.015
Individualized instruction* Student SES * limited							0.043
Individualized instruction* Student SES * mod. output-driven							0.118
Individualized instruction * Student SES * mod. input-driven							-0.018

Dependent variable: student score in reading.  
All coefficients are statistically significant at least with a  $p$ -value < 0.05 unless indicated otherwise.  
Economic development (high-income) was excluded from the regression analysis.

### ***Summary of the Findings***

In summary, multiple regression analysis confirmed a well-established relationship between student socioeconomic background and student achievement: socioeconomic background strongly predicts student achievement in reading. Out of four selected classroom practices, only individualized instruction demonstrated the potential to mitigate the predictive power of socioeconomic background. In contrast, using reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies strengthened the predictive power of the socioeconomic background variable, as they were shown to be more effective for higher SES students.

Each of the four types of teacher autonomy affected the joint influence of teacher practices and student socioeconomic background on student achievement differently. With all four classroom practices, limited teacher autonomy strengthened the predictive power of the socioeconomic background. In contrast, extended autonomy mitigated the predictive power of socioeconomic background in the context of all four classroom

practices. The implication of this is that it contributed to narrowing down the gap in student achievement. The effect of moderate input-driven autonomy varied across practices. Similar to the limited autonomy, it attenuated the influence of socioeconomic background in relation to the use of reading comprehension and autonomy-supportive strategies. In contrast, it mitigated the predictive power of socioeconomic background in the context of content differentiation and individualized instruction. Finally, output-driven autonomy strengthened the effect of socioeconomic background through the use of all classroom practices except for content differentiation.

## CHAPTER 5

### DISCUSSION AND CONCLUSION

#### **Summary of Findings**

Below I summarize the findings for all four groups of countries by the following types of teacher autonomy and corresponding types of control mechanisms over teachers: limited autonomy (i.e., strong input and strong output control mechanisms), moderate output-driven autonomy (i.e., strong input and weak output mechanisms), moderate input-driven (i.e., weak input and strong output control mechanisms) and extended (i.e., weak input and weak output control mechanisms).

#### **Group 1: Limited Teacher Autonomy**

##### ***Limited Teacher Autonomy: Country Composition***

Countries with limited teacher autonomy, i.e., with strong input and strong output control in the curriculum implementation domain, make the largest group with 19 countries. In addition to the size, it stands out as the most diverse group. Geographically, it includes countries from Western Europe (Austria, Spain), Eastern Europe (Slovenia, Georgia, Latvia, Bulgaria, Azerbaijan, Russian Federation), Middle East and North Africa (Israel, Qatar, Saudi Arabia, Morocco, United Arab Emirates), Latin America and the Caribbean (Chile, Trinidad and Tobago), North America (Canada), and East Asia (Chinese Taipei, Hong Kong, Singapore). There are significant variations in economic development as well. While the majority of the countries are high-income economies, there are upper-middle income (Azerbaijan, Bulgaria, Chinese Taipei, Russian Federation) and lower-middle income (Georgia, Morocco) economies represented as well. Countries with limited teacher autonomy also vary by the level of democratic

development: three countries are considered full democracies (Austria, Canada, and Spain), and eight countries are regarded as flawed democracies (Bulgaria, Chile, Hong Kong, Israel, Latvia, Singapore, Slovenia, Trinidad, and Tobago), two are defined as hybrid regimes (Georgia and Morocco), and six are considered as authoritarian regimes (Azerbaijan, Chinese Taipei, Qatar, Russian Federation, Saudi Arabia, United Arab Emirates).

In all these countries, teachers are subject to strong input control in the domain of curriculum implementation. In other words, their respective national or regional governments restrict their scope of action beyond developing goals and objectives of the reading instruction. Specifically, the national curriculum as a regulating document also prescribes instructional processes or methods, materials (e.g., textbooks, instructional materials), and/or assessment methods and activities. On the output side of the control, in addition to relatively softer measures, such as self-evaluation, research programs, or visits by the inspectors, the implementation of the curriculum is evaluated by national or regional examinations.

### ***Limited Teacher Autonomy: Average Student Achievement***

The results of my analysis reveal that average student achievement in the group with limited autonomy was almost as high as in those with moderate input-driven and extended autonomy. Also, the breakdown by the democratic and economic development shows that, on average, students in lower-middle-income countries with hybrid regimes, i.e., Georgia and Morocco, performed significantly lower than other countries with higher economic and democratic development levels. A slight difference in the average achievement in favor of the countries with moderate input-driven and extended autonomy

was maintained after taking important student-level (socioeconomic background, gender), school-level (availability of reading resources), teacher-level (classroom size, all four teacher qualification characteristics, and all four classroom practices), and country-level (economic and democratic development) factors into consideration.

### ***Limited Teacher Autonomy: Student Achievement Gap***

Results also indicate that the group of countries with limited teacher autonomy experiences quite a high achievement gap in reading. The difference in achievement between students in the top and bottom quartiles in the PIRLS achievement score distribution was the second highest compared to other teacher-autonomy groups. This group of countries also had the highest difference in the scores of students whose families had many and few resources. In addition, the findings showed that compared to other groups, the group of countries with limited teacher autonomy had the largest gap in the achievement of students who studied in affluent schools (the share of low SES students did not exceed 10 %) and disadvantaged schools (with more than half of their student body with disadvantaged backgrounds).

### ***Limited Teacher Autonomy: Teacher Use of Classroom Practices***

Examination of the frequency of teacher practices showed that compared to other groups, teachers in the group with limited autonomy made the most frequent use of reading comprehension and reading content differentiation strategies. In other words, teachers in this group of countries frequently engaged students in activities and tasks that target improving their reading comprehension skills. In addition, they tend to adjust the reading materials by the level and interest of their students most intensively. Educators in countries with limited teacher autonomy were also among the most frequent users of

autonomy-supportive strategies, as only those with moderate output-driven autonomy outperformed them. The findings supported the same pattern of using the strategies mentioned above after controlling for such important factors as the country's economic and democratic development level, classroom size, availability of reading resources, and four key teacher qualification characteristics: education, experience, and participation in formal and school-based professional development. Interestingly, this group also had the highest share of teachers who used individualized instruction with their students often or always.

The findings about the pattern of the frequency of using practices across schools with different shares of students with low SES may provide some explanation for existing inequities in student achievement. Compared to those who taught at schools with a lower percentage of economically disadvantaged students, teachers employed at educational institutions with more than half of their student body having low socioeconomic background used reading comprehension, content differentiation, and autonomy-supportive strategies least frequently. The use of individualized instruction was an exception, as the share of teachers applying it in the classrooms often or always was the lowest at schools with the lowest percentage of economically disadvantaged students.

Results were quite mixed when examining how teachers' perceptions of student-attribute-related limitations affecting their instruction were related to classroom behavior. On average, those teachers who thought that the effect of those factors was strong still reported using reading comprehension strategies more frequently compared to those teachers who considered these factors not affecting or somewhat affecting their instruction. The pattern was different for other practices – those teachers less concerned

about the student-attribute-related limitations over their instruction used materials considering student level and interest and gave opportunities for more student choice and critical reflection most frequently. The same group of teachers also had a higher share of those using individualized instruction in their classrooms.

### ***Limited Autonomy as a Moderator - Teacher Practices and Student Achievement***

Among the four groups by a teacher of autonomy, including limited, moderate output-driven, moderate input-driven, and extended, limited autonomy had one of the lowest (only slightly higher than extended autonomy) moderating effects on the relationship between student achievement and two classroom practices: reading comprehension strategies and use of autonomy-supportive methods. In other words, the effectiveness of the mentioned two classroom practices was relatively small in countries with limited autonomy. As findings demonstrate, the implications of limited autonomy were more desirable for using reading content differentiation and individualized instruction. Specifically, compared to other types of teacher autonomy, it had one of the strongest reinforcing effects on the positive influence of reading content differentiation. Also, its mitigating effect of the negative relationship between the use of individualized instruction and student achievement was among the strongest.

### ***Limited Autonomy – Implications for Achievement Gap***

Compared to extended autonomy, in countries with limited autonomy, using all four classroom practices was more effective for students with higher socioeconomic backgrounds. This means that limited teacher autonomy strengthened the predictive power of socioeconomic background and contributed to widening the achievement gap between students with higher and lower socioeconomic backgrounds.

### ***Limited Autonomy as a Moderator of Teacher Qualification Effect***

Limited autonomy, compared to other types of teacher autonomy, reinforced the effect of teaching experience on classroom practices least (except the content differentiation strategies). This means that experience mattered least in the countries with limited teacher autonomy for the frequency of the use of classroom practices. Compared to other forms of autonomy, it had smaller reinforcing power (similar to output-driven autonomy) on the relationship between school-based professional development and classroom activities (except for content differentiation) as well. Interestingly, compared to two other types of autonomy, limited autonomy, and moderate output-driven autonomy, more strongly reinforced the effect of both teacher qualification characteristics on the use of reading content differentiation only.

### **Group 2: Moderate Output-driven Teacher Autonomy**

#### ***Moderate Output-driven Teacher Autonomy: Country composition***

The group of countries with moderate output-driven autonomy, i.e., with strong input and weak output control in the curriculum domain, makes the smallest group, with only six countries included. Geographically, it represents countries from the Middle East (Bahrain, Oman, Iran), Central Asia (Kazakhstan), Western Europe (Portugal), and East Asia (Macao). According to the 2016 classification by the World Bank, Iran, and Kazakhstan are upper-middle-income countries, with the rest qualifying as high-income ones. Interestingly, by democratic development, all five countries, except Portugal (flawed democracy), belong to authoritarian regimes.

In all these countries, teachers are subject to strong input control in the domain of curriculum implementation. In other words, their respective national or regional



governments restrict their scope of action beyond developing goals and objectives of the reading instruction. Specifically, the national curriculum as a regulating document also prescribes instructional processes or methods, materials (e.g., textbooks, instructional materials), and/or assessment methods and activities. Teachers have more freedom with regard to the evaluation of the implementation of their curriculum. While it is evaluated by one or more of such softer mechanisms as self-evaluation, research programs, or visits by the inspectors, strong output control mechanisms, such as national and regional examinations, are not in place.

***Moderate Output-driven Teacher Autonomy: Average student achievement***

A comparison of the average student scores across the groups of countries with different types of autonomy showed that students in countries with moderate output-driven teacher autonomy had the lowest achievement. Examination of the average scores by country's economic development revealed that students in upper-middle-income countries performed better than those in the lower-middle-income. Similarly, students in countries with a higher level of democratic development (i.e., authoritarian regimes vs. flawed democracy) demonstrated better performance. The lowest average performance of the students in the countries with output-driven autonomy was also confirmed by the regression analysis after controlling for student-level (socioeconomic background, gender), school-level (availability of reading resources), teacher-level (classroom size, all four teacher qualification characteristics, and all four classroom practices), and country-level (economic and democratic development) factors.

***Moderate Output-driven Teacher Autonomy: Student Achievement Gap***

Results also indicate that the group of countries with moderate output-driven teacher autonomy experiences quite a high achievement gap in reading. The difference in achievement between the top and bottom quartiles in the PIRLS achievement score distribution was the highest compared to other teacher-autonomy groups. Also, the discrepancy in the scores was the second highest when examining equity-related statistics by the differences in achievement between students whose families had many and few resources. In addition, the findings demonstrated the second largest gap in the achievement of students who studied in the most affluent schools, i.e., the share of low SES students did not exceed 10 %, and the poorest schools, which had more than half of their student body with disadvantaged backgrounds.

***Moderate Output-driven Teacher Autonomy: Teacher Use of Classroom Practices***

As results indicate, compared to other autonomy types, teachers in the countries with moderate output-driven autonomy used the autonomy-supportive strategies most frequently. They also made the second most frequent use of reading content differentiation. In addition, less than half of the teachers, yet a higher share than in moderate input-driven and extended autonomy, reported using individualized instruction often or always. Finally, compared to other strategies, teachers in this group made relatively less frequent use of content differentiation, teachers with limited and moderate input-driven autonomy outperforming them. The findings supported the same pattern of using reading comprehension, content differentiation, and autonomy-supportive strategies after controlling for such important factors country's economic and democratic development level, classroom size, availability of reading resources, and four key teacher

qualification characteristics: education, experience, and participation in formal and school-based professional development.

Similar to limited autonomy, in countries with moderate output-driven autonomy, teachers at schools with a higher share of disadvantaged students used reading comprehension, content differentiation, and autonomy-supportive strategies less frequently. Also, for the most part, the higher the percentage of disadvantaged students was at school, the proportionally fewer teachers used individualized instruction.

I found a similar pattern with regard to teachers' perception of their instruction being limited by student-related attributes. The results indicated the highest frequency of using all four strategies among those teachers who considered that student-related attributes affected their instruction only to a small extent.

### ***Moderate Output-driven Teacher Autonomy as a Moderator - Teacher Practices and Student Achievement***

Compared to other forms of teacher autonomy, a moderate output-driven one showed the strongest reinforcing power for the effectiveness of the following three strategies: reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies. Thus, after accounting for student-level, school-level, teacher-level, and country-level factors, the same frequency of all four classroom practices yielded the highest student scores in the countries with moderate output-driven teacher autonomy. It should be mentioned that, at the same time, moderate output-driven teacher autonomy most strongly reinforced the negative effect of individualized instruction on student achievement.

### ***Moderate Output-driven Teacher Autonomy: Implications for Achievement Gap***

Compared to extended autonomy, moderate output-driven autonomy strengthened the effect of socioeconomic background by using all classroom practices except for content differentiation. This means that reading comprehension strategies, autonomy-supportive strategies, and individualized instruction were more effective for students with higher socioeconomic backgrounds. Thus, for the most part, moderate output-driven autonomy contributed to increasing discrepancies between the achievement of higher and lower SES students.

### ***Output-driven Teacher Autonomy as a Moderator of Teacher Qualification Effect***

Compared to other types of teacher autonomy, moderate output-driven autonomy reinforced the effect of teaching experience on using all three practices with the highest magnitude. In other words, compared to other countries, for the frequency of classroom practices, teaching experience mattered the most. The findings were different in relation to another teacher qualification – school-based professional development. Compared to other forms of autonomy, it had smaller reinforcing power (similar to the limited autonomy) on the relationship between school-based professional development and classroom activities (except for content differentiation).

### **Group 3: Moderate Input-driven Teacher Autonomy**

#### ***Moderate Input-driven Teacher Autonomy: Country composition***

Countries with moderate input-driven freedom from control, i.e., weak input and strong output control mechanisms, make the second largest group among the four. Despite its large size (11 countries), it is quite homogenous in terms of geographic representation and economic and democratic development. All countries represent Western Europe (Belgium (French), Denmark, France, Germany, Ireland, Italy, the

Netherlands, Northern Ireland, and Sweden)) and Eastern Europe (Slovak Republic), with the addition of Australia from the Pacific. In terms of economic development, all of them are high-income economies. Finally, all the countries have advanced democratic development and qualify as full or flawed democracies.

In all these countries, teachers are subject to weak input control. This means that the national curriculum only defines the goals and objectives of the reading instruction. Teachers are free to decide what instructional or assessment methods and materials (such as textbooks) to use in their classrooms. However, for teachers in these countries, the freedom on the input side is constrained by the strong output control mechanisms. Specifically, how they implement the curricula is evaluated not only through school self-evaluations, research programs, or visiting inspectors but also by national or regional examinations.

***Moderate Input-driven Teacher Autonomy: Average Student Achievement***

Students in countries with input-driven autonomy were among the highest achievers. Their average achievement score in reading was slightly lower than that of the students from the countries with extended autonomy. The results indicate a slight difference in the average achievement by democratic development level in favor of full democracies. Even after controlling for student-level (socioeconomic background, gender), school-level (availability of reading resources), teacher-level (classroom size, all four teacher qualification characteristics, and all four classroom practices), and country-level (economic and democratic development) factors, the average score in these groups of countries was the second highest.

***Moderate Input-driven Teacher Autonomy: Student Achievement Gap***

In addition to high average student achievement, countries with moderate input-driven autonomy (together with the extended ones) had the smallest achievement gap in reading. The difference between students in the top and bottom quartile in the PIRSL 2016 distribution was slightly higher than that of countries in extended autonomy. The discrepancy in the scores was the lowest of all groups when examining equity-related statistics by the differences in achievement between students with high and low socioeconomic backgrounds, as well as between the schools with the lowest and highest share of economically disadvantaged students.

***Moderate Input-driven Teacher Autonomy: Teacher Use of Classroom Practices***

Interestingly, teachers with input-driven autonomy, together with those with extended autonomy, used reading comprehension and autonomy-supportive strategies least frequently. As results indicate, unlike those two classroom practices, teachers in this group of countries used reading content differentiation most frequently, together with those with limited autonomy. The regression results supported the same relative frequency of using reading comprehension, content differentiation, and autonomy-supportive strategies after controlling for such important factors country's economic and democratic development level, classroom size, availability of reading resources, and four key teacher qualification characteristics: education, experience, and participation in formal and school-based professional development. Finally, compared to all other countries, the lowest share of teachers with moderate input-driven autonomy used individualized instruction often or always.

Compared to other groups of countries, the differences in the frequency of the use of classroom practices across the schools with different shares of low SES students were

the lowest in counties with moderate input-driven autonomy. Yet, in most cases, the use of these practices favored schools with a higher share of disadvantaged students (e.g., reading comprehension strategies, reading content differentiation, and individualized instruction).

Results were quite mixed when examining how teachers' perceptions of student-attribute-related limitations affecting their instruction were related to classroom behavior. On average, those teachers who thought that the effect of those factors was strong reported using reading content differentiation strategies more frequently than compared to those teachers who considered these factors not affecting or somewhat affecting their instruction. The pattern was different for other practices – those teachers less concerned about the student-attribute-related limitations over their instruction used reading comprehension and autonomy-supportive strategies slightly, yet, more frequently. Finally, the difference in the share of teachers more and less regularly using individualized instruction was found to be very small across schools with different percentages of economically disadvantaged students.

### ***Moderate Input-driven Autonomy as a Moderator - Teacher Practices and Student Achievement***

Compared to other forms of autonomy, a moderate input-driven one showed one of the strongest reinforcing power for the effectiveness of reading comprehension and autonomy-supportive strategies. Thus, after accounting for student-level, school-level, teacher-level, and country-level factors, the same frequency of all four classroom practices yielded one of the highest student scores in the countries with moderate input-driven teacher autonomy (only those in countries with output-driven autonomy

outperformed them). Moderate input-driven autonomy (similar to extended autonomy) had the smallest moderating effect on the use of reading comprehension strategies. Also, its mitigating impact on the negative relationship between the use of Individualized instruction and student achievement was among the strongest.

#### ***Moderate Input-driven Teacher Autonomy: Implications for Achievement Gap***

The moderating effect of input-driven autonomy on the relationship between student socioeconomic background and achievement varied across different practices. Compared to the extended autonomy, it attenuated the influence of socioeconomic background in relation to the use of reading comprehension and autonomy-supportive strategies. In contrast, similar to extended autonomy, it mitigated the predictive power of socioeconomic background in the context of content differentiation and individualized instruction.

#### ***Moderate Input-driven Teacher Autonomy as a Moderator of Teacher Qualification***

Moderate input-driven autonomy, together with the output-driven one, most strongly and positively moderated the effect of teaching experience on the use of reading comprehension strategies and autonomy-supportive strategies. Its moderating effect was smaller in relation to content differentiation. Similarly, moderate input-driven autonomy most substantially strengthened the positive impact of school-based professional development on the use of reading comprehension strategies and autonomy-supportive strategies. Finally, compared to other types of autonomy, its moderating effect was one of the lowest for reading content differentiation.

#### **Group 4: Extended Teacher Autonomy**

##### ***Extended Teacher Autonomy: Country composition***



The countries with extended autonomy, i.e., weak input and weak output control mechanisms, make another homogenous cluster. Geographically, it resembles group 3: Western Europe (Belgium (Flemish), Finland, Norway), Eastern Europe (Hungary, Czech Republic, Lithuania), and the Pacific (New Zealand). All the countries are advanced in terms of both economic and democratic development. By 2016, when PIRLS 2016 was administered, all qualified as high-income countries. By the Democracy Index, they ranked as full democracies (Finland, New Zealand, Norway) or flawed democracies (Belgium, Czech Republic, Hungary, Lithuania).

Teachers in these seven countries are granted freedom to the greatest extent as both input and output control they are subject to is weak. This means that the national curriculum only defines the goals and objectives of the reading instruction. Teachers are free to decide what instructional or assessment methods and materials (such as textbooks) to use in their classrooms. They also have more freedom with regard to the evaluation of the implementation of their curriculum. While it is evaluated by one or more of such softer mechanisms as self-evaluation, research programs, or visits by the inspectors, strong output control mechanisms, such as national and regional examinations, are not in place.

### ***Extended Teacher Autonomy: Average Student Achievement***

Results of this study indicated that, on average, students in countries with extended teacher autonomy had the highest reading scores. There was only a small variation in the average achievement by democratic development level in favor of full democracies. The highest average performance of students in the countries with extended autonomy was also confirmed by the regression analysis after controlling for student-

level (socioeconomic background, gender), school-level (availability of reading resources), teacher-level (classroom size, all four teacher qualification characteristics, and all four classroom practices), and country-level (economic and democratic development) factors.

### ***Extended Teacher Autonomy: Student Achievement Gap***

In addition to high student achievement, countries with extended autonomy - similar to the moderate input-driven ones - had the smallest achievement gap. Specifically, they had the lowest difference between the reading scores of the students in the top and bottom quartile in the PIRSL 2016 distribution. The difference in the scores was the second lowest of all groups when examining equity-related statistics by the differences in achievement between students with high and low socioeconomic backgrounds, as well as between the schools with the lowest and highest share of economically disadvantaged students.

### ***Extended Teacher Autonomy: Teacher Use of Classroom Practices***

Interestingly, the results indicated that both before and after controlling for important countries, school, and teacher level factors, teachers with extended autonomy used reading comprehension, content differentiation, and autonomy supportive- strategies least frequently (before controlling for the mentioned variables, they only outperformed teachers with moderate input-driven autonomy in the use of reading comprehension strategies). Also, these countries had one of the lowest shares of teachers using individualized instruction often or always.

Examination of teacher practices by the share of low SES students at schools revealed that overall, in schools with a higher share of disadvantaged students, teachers

used all classroom practices more frequently. Similarly, teachers who thought that student attributes limited their instruction used these practices more frequently.

### ***Extended Teacher Autonomy as a Moderator - Teacher Practices and Student Achievement***

Compared to other forms of autonomy, extended autonomy had the lowest reinforcing effect on the relationship between student achievement and (1) comprehension strategies, (2) reading content differentiation, and (3) use of autonomy-supportive strategies. Finally, compared to limited and moderate input-driven autonomy, individualized instruction was more effective in the countries with extended autonomy.

### ***Extended Teacher Autonomy: Implications for Achievement Gap***

In the countries with extended teacher autonomy, teachers' use of all four classroom practices was more effective for students with lower SES. This means that extended autonomy demonstrated the potential for mitigating the predictive power of socioeconomic backgrounds and narrowing the gap in student achievement.

### ***Extended Autonomy as a Moderator of Teacher Qualification Effect***

Extended autonomy had the lowest or one of the lowest moderating effects on the relationship between teaching experience and the use of all three classroom practices. In other words, compared to other types of autonomy, teaching experience had a weak association with classroom practices in the countries with extended autonomy. The findings were mixed in relation to school-based professional development. Specifically, after the moderate input-driven autonomy, it most strongly reinforced the effect of school-based activities on using autonomy-supportive and reading comprehension strategies. In contrast, its impact was the lowest for reading content differentiation.

## **Discussion and Implications**

### **Distribution of PIRLS 2016 Countries by Type of Teacher Autonomy**

As discussed in the second chapter, introducing neoliberal and market-driven principles since the 1980s has had substantial implications for the national educational systems. Emphasis on marketization, managerialism, and competitive performativity had a significant effect on the teaching profession (Gobby et al., 2018), turning them into skilled technicians, or managed professionals, with narrowly described technical competencies and the obligation of producing pre-determined outcomes (Codd, 2005). At the same time, as the scholars point out, despite the vast global influence of neoliberal principles, there were variations in the design and implementation of specific educational-level policies across the countries (Davies & Bansel, 2007; Silova, 2018; Takala & Piattoeva, 2012).

Analysis of the PIRLS 2016 data revealed existing incoherencies and differences in the design and implementation of the curriculum-related policies, which were one of the major reforms areas of neoliberal and market-driven reforms. It also sheds light on more recent developments in the country's policy-level choices with regard to standardizing curriculum, i.e., creating input-based control mechanisms, and institutionalizing performance evaluation, i.e., creating output-based control mechanisms. As the distribution of the PIRLS 2016 countries by the types of autonomy demonstrates, despite the initial influence of neoliberal reforms on educational policies, high-income countries with advanced democratic development from Europe and the Pacific region and Oceania mostly institutionalize weak input national-level control mechanisms for curriculum implementation paired with weak or strong output control mechanisms. Such

a composition of countries with extended or moderate input-driven teacher autonomy, as well as the depicted differences in the selection of the type of output control, can be explained by the political and socio-cultural factors and the historical commitment of some of these countries to school and teacher autonomy. For example, as the findings indicate, all Nordic countries with a strong commitment to welfare state principles and historically lower degree of control compared to Anglo-American counterparts (Mausethagen & Molstad, 2015; Stephens et al., 2004) were found to be among those favoring weak input control. In addition, Finland, with the cultural goal of schooling and equality (Salonen-Hakomaki et al., 2016) and historical adherence to teacher autonomy (Paradis et al., 2018; Sahlberg, 2011), is shown to favor weak output control, i.e., absence of high-stakes national examinations. On the other hand, Denmark, which has been observed to favor stronger accountability mechanisms compared to other Nordic countries (Paulsrud & Wermke, 2020), is among the countries that grant input-driven autonomy. My findings also showed that Norway, although considered a relatively strong supporter of national-level control and lower teacher autonomy (Bakken, 2019, Carlgren & Klette, 2008), was among those with the highest level of teacher autonomy, i.e., extended one.

The findings of this study were aligned with the scholarly analysis of the degree of historical commitment to the principles of neoliberalism and a market-driven economy. A higher share of the countries allowing for weak input control still maintained strong accountability mechanisms in place by instituting national or regional examinations as one of the mechanisms for evaluating the curriculum implementation. Among those most dedicated to those principles, New Zealand makes an exclusion. Like

Finland and Norway, they are among those allowing for extended autonomy to teachers with minimal accountability mechanisms.

The political and socio-cultural factors I discussed in Chapter 2 help explain the composition of the groups of countries with strong input control, i.e., limited and moderate output-driven teacher autonomy. For example, the presence of Arab-speaking and other Asian countries (also with authoritarian regimes) was an expected finding. Despite declared efforts to learn lessons from the Western countries, deeply embedded socio-cultural factors and the historical absence of teacher professional freedom and agency (Chaaban et al., 2021; Mustafa & Cullingford, 2008), these countries are widely represented in both groups by teacher autonomy. The governments' quest for exceptional economic competitiveness via robust educational systems has resulted in strong performativity culture in East-Asian countries, explaining the restricted nature of teacher autonomy. In addition, the legacy of extreme centralization and a culture of compliance and conformity sheds light on the presence of post-Soviet countries among those with limited and moderate output-driven autonomy (Khachatryan et al., 2013; Teleshalyev, 2013). Finally, as the list of the countries in the findings section shows, a smaller number of European countries, as well as Canada, have stayed committed to market-driven rigid control mechanisms. In the case of Canada, this can also be explained by the historical distrust towards the teachers both from the government and the society (Paradis et al., 2018).

### **Teacher Qualification and the Use of Classroom Practices**

Before exploring the moderating power of teacher autonomy on the effect of two important teacher qualification characteristics, teaching experience and school-based

professional development, I examined the relationship between each of the two and teacher use of three classroom practices: reading comprehension strategies, reading content differentiation, and individualized instruction. My findings indicated that when holding country, school, and teacher level factors constant, including teacher education and intensity of participation in formal and school-based professional development, teaching experience predicted the frequency of the use of these strategies. In other words, the more years of experience teachers had, the more frequently they used these strategies in the classroom. This finding contradicted the one by Guo et al. (2012), who found a negative association between time on academic activities, in general, and the number of teaching years. However, it is aligned with the results of a recent review of US-based studies (Podolsky et al., 2019), according to which teaching experience, although depending on what context is being accumulated in, is positively associated with teacher effectiveness, also implies the use of effective strategies in the classroom.

It should be noted that among all three strategies, including reading content differentiation and autonomy-supportive strategies, the experience was the strongest predictor, with the largest effect size, for using reading comprehension strategies. In contrast, it predicted the use of autonomy-supportive strategies with the smallest magnitude and also explained the smallest variation in teacher use of these practices. This result about autonomy-supportive strategies did not support earlier evidence of the predictive power of teaching experience being exceptionally high for autonomy-supportive practices as they increase confidence and self-efficacy among educators resulting in releasing a significant amount of control to the students (Huberman, 1992; Klassen & Chiu, 2010). This difference in the findings may be partially explained

by the fact that the cited authors examined the use of autonomy-supportive strategies in light of classroom management, while my research focused specifically on using these strategies in the process of reading comprehension.

This study yielded interesting findings about the relationship between teachers' participation in school-based professional development and the use of classroom practices. Compared to teaching experience, it had far stronger predictive power for all classroom practices and also explained the higher variation in teacher actions. Thus, this study demonstrated that the more frequently teachers shared their learning with colleagues, observed peers' lessons, and worked together to teach specific topics, the more frequently they used all three classroom practices. These findings are well aligned with the socio-cultural perspective on learning, according to which collaborative work enables teachers not only to expand their knowledge and understanding of effective strategies but, most importantly, helps them to internalize new learnings and make part of the everyday practice (Levine & Marcus, 2007). Yet, as there are studies that indicate differential effects of school-based professional development activities on teacher practices and, eventually, student achievement (Guskey & Sparks, 2004; Ke & Huang, 2019), this finding contributes to the existing body of literature by identifying specific types of school-based professional development that contribute to improved classroom practices. At the same time, when interpreting these results, one should be aware of the purely quantitative nature of the measure of teacher practices, which does not capture the quality dimension of the key three classroom practices. As Ke and Huang (2019) point out, this is a crucial point to consider. Their findings revealed that such school-based activities as collective lesson planning and teacher collegiality, in some cases, improved



the frequency of the use of desirable teaching strategies but not necessarily the quality of delivery.

The findings of this study related to the effect of the characteristics of teacher qualifications on the frequency of teacher classroom practices have strong implications for national or school-level policies. Given such an exceptionally solid influence of school-based professional development experience on teacher behavior, governments and schools should reorganize limited financial or human resources and organize the teaching and learning process in such a way that teachers are given more opportunities to share their experiences, observe each other's classroom, work together on specific educational topics and also collaborate across different grade levels. Due to continuously evolving methodological knowledge in the field of education, creating such support systems can help teachers contextualize new knowledge, expand the newly desired skillset and successfully respond to changing education landscape for the best interests of their students.

### **Effect of Teacher Autonomy on Teacher Classroom Practices**

Existing literature quite intensively discusses the question of the suitability of high-degree autonomy for all teachers and raises questions about its universal benefit. As researchers argue, it is risky to grant teachers autonomy when they do not possess sufficient mastery and professional maturity (Anderson, 1987; Ashley, 2012; Erss, 2018; Grant et al., 2020; Littlewood, 1996; Strong & Yoshida, 2014; White, 1992). At the beginning of their career, their self-governance capacity is lower and does not allow them to handle freedom from control properly and translate it into effective classroom behavior (Fox, 2012; Grant et al., 2020; Gwaltney, 2012). To create additional empirical evidence

in relation to this teacher competence/qualification-related concern, I examined if different types of country-level teacher autonomy in the domain of curriculum implementation moderated the relationship between teacher qualification characteristics and the frequency of the use of classroom practices differently. In other words, the purpose of this inquiry was to understand if extended and moderate input-driven autonomy, both requiring a lot of independent and self-initiated actions, more strongly and positively moderated the relationship between teacher qualifications and the frequency of classroom practices compared to the limited and moderate output-driven types of autonomy.

As described in the fourth chapter, I tested the moderating effect of teacher autonomy on the impact of two teacher qualification characteristics (teaching experience and school-based professional development) on the following classroom practices: reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies. First of all, as I already explained above, teaching experience and school-based professional development were positively and significantly associated with all three classroom practices. In other words, the more years of teaching experience teachers had, and the more intensively they participated in school-based professional development, the more frequently they used all three classroom practices with their students.

As the results indicate, the moderating effect of the four types of teacher autonomy varied for each of the teacher qualification characteristics. Among the four types of autonomy, the moderate output-driven one most strongly moderated the relationship between teaching experience and three classroom practices. Input-driven

autonomy was the second strongest moderator. This means that in countries with limited and extended autonomy, the number of teaching experiences had a relatively small importance for the frequency of classroom practices.

The moderating pattern of teacher autonomy was different for school-based professional development. These findings suggest that limited and moderate output-driven autonomy strongly reinforced the relationship between school-based professional development and reading content differentiation. At the same time, they had the smallest moderating effect in relation to reading comprehension and autonomy-supportive strategies. Compared to others, moderate input-driven autonomy consistently strengthened the effectiveness of school-based professional development for all three classroom practices.

Thus, my findings partially support the claims in the scholarly literature, according to which when teachers are granted high-level autonomy in deciding what and how to teach (i.e., freedom from input control), their classroom behavior will significantly depend on their qualifications. As shown, teaching experience mattered least for teachers with the highest degree of autonomy (i.e., extended autonomy). This means that even though they had the freedom to decide about the nature of the instructional process rather than use the plans and resources prescribed by the curriculum, the frequency of their use of practices, compared to the teachers with other types of autonomy, was determined less by the number of their teaching years. Compared to teaching experience, overall, participation in school-based professional development activities mattered more for teachers with extended (as well as moderate input-driven autonomy) compared to limited and moderate output-driven. This may create a concern

that these teachers may be less effective in their classrooms without solid and relevant professional development. On the other hand, the relatively higher moderating power of school-based professional development in countries with extended autonomy, compared to the limited and moderate output-driven one, may indicate high-level relevance and meaningfulness of those activities in the former.

Overall, the finding about the teachers with moderate input-driven autonomy is more consistently supportive of the propositions of the framework of “graduated autonomy” (Fox, 2012). Both teaching experience and intensity of school-based professional development mattered more for these teachers compared to those with extended and, most importantly, limited autonomy. This difference may be related to the presence of solid output control. However, future research is needed to understand if and why the parallel presence of strong output control mechanisms makes teachers' classroom behavior more conditional to their qualifications in countries with moderate input-driven autonomy compared to the extended one.

### **Effects of Classroom Practices on Student Achievement**

As part of my analysis, I examined the relationship between teacher use of key classroom practices and student achievement. As the findings suggest, three practices – reading comprehension strategies, reading content differentiation, and autonomy-supportive strategies – were positively associated with student achievement. This result is consistent with other researchers' findings that teacher classroom instruction and classroom support, in general, positively affect student achievement (Guo et al., 2012). Among them, reading comprehension strategies, i.e., having students engage in such activities as identifying main ideas, explaining and supporting what they read, comparing,

making predictions, and generalizations, had the strongest predictive power for student achievement. This finding is well aligned with existing scholarly literature, which also documents that explicit instruction of reading comprehension strategies in the primary grades improves student achievement in reading (Amin, 2019; Dwiningtiyas et al., 2020; Küçükoğlu, 2013). As Paris and Paris (2007) explain, teachers' intentional and targeted support for developing reading comprehension skills is key as the latter is a very complex process requiring a lot of effort students cannot undertake independently.

The impact of reading content differentiation was of a smaller magnitude yet a significant predictor of student achievement in reading. As the selection of the materials by students' reading level and interest is widely believed to be an important aspect of responding to specific student needs (Heacox, 2018; Tomlinson, 2014), this finding also confirmed the power of this important aspect of differentiated instruction.

Compared to the other two classroom practices, the teacher's use of autonomy-supportive strategies in the classroom, specifically allowing to choose reading texts, challenge opinions, develop their understanding of the text, or give them individualized feedback, had a smaller positive effect on student achievement in reading. Compared to reading comprehension strategies and content differentiation, approaches in the scholarly literature to the benefits of using student autonomy-supportive strategies are more contradictory. On the one hand, it is believed to foster motivation, engagement, and self-regulation in learning (Alley, 2019; Evans & Boucher, 2015). At the same time, implementing effectively in certain national cultures is considered less appropriate or more challenging (Boyadzhieva, 2016). Although this finding does not allow to specify the mechanisms through which those autonomy-supportive strategies improved student

achievement, it creates additional evidence supporting the scholarly literature, which argues that learner autonomy enhances student learning and academic performance.

Unlike other classroom practices, results indicated a negative relationship between the frequent use of individualized instruction and student achievement in reading. Students of those teachers who often or always used this practice scored lower than those whose teachers never or only sometimes used individualized instruction. According to the scholarly and methodological literature, individualized instruction, as one of the strategies for differentiated instruction, is expected to improve student achievement, as by design, it “allows all students to access the same classroom curriculum by providing entry points, learning tasks, and outcomes tailored to students’ learning needs” (Watts-Taffe et al., 2012, p. 304). Also, the absence or lack of individualized approaches and more emphasis on whole-class instruction harms students with learning disabilities the most (Schumm et al., 2000). The contradictory finding of my study may indicate that while working individually with selected students, teachers may unintentionally remove necessary support time for other students.

### **The Implications of Teacher Autonomy for Student Achievement**

Examination of the relationship between classroom practices and student achievement showed that three practices, using reading comprehension strategies, content differentiation, and autonomy-supportive strategies, were positively associated with student achievement. I expected that results would indicate a higher use of these practices in countries with moderate input-driven and extended autonomy for the following three reasons. First, as both descriptive and inferential statistical analysis with controlling for important factors revealed, students in these two groups of countries performed higher

than those with limited and moderate output-driven teacher autonomy. Second, as argued in the literature, teachers who have the freedom to decide how and what to teach their students (i.e., have freedom from input control) better tailor their instruction to students' needs and use individualized approaches more frequently (LaCoe, 2008, Prichard & Moore, 2016; Salokangas & Ainscow, 2017; Skaalvik & Skaalvik, 2009). Third, existing research also suggests that a higher degree of teacher autonomy has the potential to enhance and develop student autonomy to a greater extent (Benson, 2011; Leroy et al., 2007). My findings did not align with these arguments as, on average, teachers with extended and moderate input-driven autonomy used reading comprehension strategies, reading content differentiation, and even autonomy-supportive strategies less frequently than their colleagues in countries with limited and moderate output-driven autonomy.

I also tested the relative effectiveness of classroom practices across countries with different types of autonomy. Specifically, I was interested in examining if teachers with extended autonomy and input-driven autonomy, who, as Webb (2002) would argue, were free from the requirement of adhering to teacher proof-curricula and instead could use their professional judgment extensively, were able to achieve the better student performance in reading through the use of above-mentioned three classroom practices. The moderation analysis revealed that compared to other forms of autonomy, a moderate output-driven one showed the strongest reinforcing power for the effectiveness of all three strategies. Teachers' use of reading comprehension and autonomy-supportive strategies were also quite effective, i.e., yielding high student scores, in the countries with moderate input-driven autonomy. Compared to others, limited and extended autonomy had the slightest effect. Among the two, using the three classroom practices was less

effective in the counties with extended autonomy. In other words, with the same frequency of the three strategies, students in countries with extended autonomy scored the lowest.

The relatively small effectiveness of key instructional practices in the hands of teachers with limited autonomy could be explained by the rigid control mechanisms teachers experience, limiting their flexibility to respond to student needs and minimizing the one-size-fits-all approaches of highly prescriptive curriculum (Olivant, 2015). The same line of thinking would fail to explain even less effectiveness of classroom practices in the hands of teachers with extended autonomy. Opponents of teacher freedom would justify such a low moderating effect of the extended autonomy by the risks associated with teacher freedom, such as opportunistic behavior (Woessmann, 2005) or the temptation to maintain the status quo and follow the old patterns teachers are used to without any critical evaluation (Erss, 2018; Johnson, 2016).

One of the noteworthy findings of this analysis is the difference in the effectiveness of teacher-classroom practices between the systems with moderate input-driven and extended autonomy. As I described above, there is a lot of similarity between the countries across these two groups by economic and democratic development, as well as geographic location. Given these similarities, it is worth examining if one of the major differences, the presence of strong output control in the form of high-stakes regional examinations, accounts for the relatively higher effectiveness of teacher practices as the proponents of accountability measures (Hamilton et al., 2002; Hanushek, 2019) would also argue.

### **Student Socioeconomic Background and Student Achievement**



As hypothesized, my study revealed a well-established relationship between student socioeconomic background and student achievement (Chmielewski, 2019; Hattie, 2009; Sandsør et al., 2023). The amount of such home resources as parental education, parental occupation, home study support, and the number of books at home strongly predicted student achievement in reading. After controlling for all student, school, and country-level variables, student socio-economic background still significantly impacted student reading scores. Its effect was twice as much as those of all country-level, school, teacher, and student-level predictors combined.

### **Effects of Classroom Practices on Achievement Gap**

To understand the role of teacher autonomy in mitigating the effect of student socioeconomic background, i.e., closing the achievement gap, I first examined how the four selected classroom practices moderated the relationship between student socioeconomic background and student achievement. The results indicated that none of the three classroom practices, showing a positive association with student achievement, could mitigate the effect of socioeconomic background on student achievement. Moreover, they positively moderated the relationship between the two variables, leading to a larger gap in reading achievement. In other words, the higher students' socioeconomic background was, and the more frequently the teacher used classroom practices, the higher students scored.

The pattern was different for the individualized instruction. Although it showed a negative association with student achievement, it did slightly contribute to reducing an SES-based discrepancy in reading scores. Thus, except for the latter, three classroom practices disproportionately benefited students and did not show the capacity to enable

lower SES students to keep up with their peers from families with more resources. Few studies have examined the similar role of classroom practices in mitigating or exacerbating the achievement gap. The findings were not conclusive. For example, Atlay and colleagues (2019) found that classroom management strategies used by teachers enhanced the learning of students with high socioeconomic backgrounds and could not compensate for low socioeconomic backgrounds. Another study by Stienstra et al. (2022) found a compensatory effect of classroom practices on student achievement as their positive influence was stronger for lower SES students. Mostly, my findings generated supporting evidence for the differential effects of these practices on students with high and low SES. However, it also shows that some practices, such as individualized instruction, may still compensate for the socioeconomic background.

### **Implications of Teacher Autonomy for Achievement Gap**

One of the major concerns about granting a high degree of autonomy to teachers is related to the potential risk of exacerbating inequities among advantaged and less advantaged students (LaCoe, 2008; Murphy et al., 1986). This body of literature fears that instead of taking advantage of the freedom to tailor instruction to their students' needs, teachers may deny access to rigorous curriculum to some students based on their perceptions of these students' abilities (LaCoe, 2006). One of the most recent studies by Cimentada (2020) also revealed the differential effect high-level autonomy may have on the bottom and top achievers in certain subjects.

To understand how teacher autonomy, as a freedom from the national level control in the curriculum implementation, is related to the student achievement gap, I first examined patterns of teachers using different classroom practices across different types

of schools. There is evidence that teachers at schools with different SES compositions behave differently. For example, those at high-poverty schools trust their students less and use autonomy and constructive approaches less frequently (Solomon et al., 1996). As the school composition by student socioeconomic background has a strong association with student achievement (Rumberger & Palardy, 2005; Sirin, 2005), understanding the pattern of teacher behavior in high-poverty schools is important across countries with different types of autonomy can be helpful to explain the size of achievement gaps and also identify the potential of each to mitigate the effects of socioeconomic background.

The findings of my study indicated that teacher use of classroom practices across schools with different SES compositions varied by the type of teacher autonomy. Overall, in countries with moderate output-driven and limited autonomy, teachers at schools with a higher share of low SES students used all four classroom strategies less frequently than those engaged with more affluent students and schools. It should be noted that the magnitude of differences was exceptionally high in the countries with moderate output-driven teacher autonomy. In systems with extended autonomy, teachers at schools with a higher share of low SES students used these classroom practices more frequently. Mainly, a similar pattern was observed in countries with moderate input-driven autonomy. Along the same lines, my findings showed that in countries with extended autonomy, teachers, who thought that student attributes limited their instruction, still used classroom practices more frequently than their peers, less concerned by the same limitations. The pattern was reversed or with smaller variations in other groups of the countries.

To explore the implications of each type of autonomy for the achievement gap, I tested the secondary moderating effect of autonomy on the influence of teacher-classroom practices on the relationship between students' socioeconomic background and student achievement. Among the limited, moderate output-driven, moderate input-driven, and extended autonomy, only an extended one showed consistent potential to mitigate the strong effect of socioeconomic background often reinforced by teacher classroom practices. Moderate input-driven autonomy showed less consistent, although more potential also to contribute to narrowing achievement gaps, compared to limited and moderate output-driven forms of autonomy.

Thus, findings about the behavior patterns of teachers with different types of autonomy, as well as the relative potential of different types of autonomy to mitigate the strong predictive power of socioeconomic background, contradict the concerns in the literature about the risks to equity related to high degree teacher autonomy (LaCoe, 2008; Murphy et al., 1986). First, unlike limited and moderate output-driven autonomy, teachers in countries with extended and moderate input-driven autonomy more frequently used their classroom practices with those students who may needed them most, i.e., students with low socioeconomic backgrounds. Also, it can be implied from the findings that when teachers with extended and moderate input-driven autonomy saw that students in their class lacked prerequisite knowledge or skills, experienced a lack of sleep or food, or were distracted and not dedicated to learning, they tried to compensate for those disadvantages by more frequent use of effective classroom practices. Findings also showed that teachers' practices in countries with extended autonomy contributed to narrowing down the gap. Although, overall, this effect was not similarly consistent in the

countries with input-driven autonomy, they still showed more potential to benefit disadvantaged students compared to the ones with limited and moderate output-driver autonomy. Such attitudes and behavior patterns also explain significantly lower achievement gaps in the countries with weak input control, i.e., moderate input-driven and extended autonomy.

Finally, comparing average student achievement and the size of the achievement gap in reading between the countries with extended and moderate input-driven teacher autonomy allows me to draw conclusions about the educational performance and equity effects of using accountability. First, the implications of the findings of this study are aligned with the research, according to which standardized and high-stakes testing has no relation to the performance of students in reading (Nichols et al., 2006) and is not a much-needed intervention for improving educational outcomes (Sahlberg, 2010). Per the findings of my study, on average, students in countries where teacher autonomy was limited by strong output control mechanisms in the form of national or regional examinations (i.e., moderate input-driven teacher autonomy) did not perform better than those in the countries where accountability measures were absent but otherwise were quite similar (i.e., extended teacher autonomy).

Equity-related implications of this study also question the suitability of accountability measures. Scholarly literature supporting such efforts argues that national educational leaderships have failed to create more equitable systems in the absence of strong accountability measures and the evident inability to minimize the negative effects of teacher prejudice and low expectations (Scheurich & Skrla, 2004). The findings of this study contradict such an understanding of the importance of accountability. The presence

of accountability measures, i.e., the mechanisms of output control, was one of the major differences between the economically and democratically similar countries with extended and moderate input-driven teacher autonomy. As discussed, the findings indicated that overall, the size of the achievement gap as well as teachers' use of educational practices in favor of disadvantaged students, was similar in these two groups of countries. In addition, classroom practices were more effective for lower SES students in countries with extended teacher autonomy. Thus, I can argue that, at minimum, in the context of the countries with a high degree of autonomy on the input side, the presence of accountability mechanisms did not have positive implications for student equity.

### **Study Limitations**

My study is subject to several limitations. First, given its design, instead of demonstrating causal connections between the types of teacher autonomy and other variables of the main interest, it uses a comparative inquiry to make suggestions and highlight implications of different forms of input- and output-control-based autonomy. Second, this study has limited generalizability as it analyzes the data of those countries that participated in PIRLS 2016. In addition, due to the missing sections of the data, I had to exclude additional countries from the existing PIRLS sample. Related to this, not all four groups of countries by teacher autonomy include countries with all levels of economic and democratic development. This has limited the within-group comparison of countries with different levels of economic and democratic development. In other words, I was not able to assess the intersectionality of autonomy, democratic and economic development in relation to teacher practices, qualifications, and student achievement. Another limitation is related to the source of teacher-related data. Given cultural

differences across countries, the accuracy of the teacher-reported data may vary and not provide an equally objective representation of their classroom contexts.

### **Conclusion**

Despite globalization and economy-driven neoliberal influences on educational systems worldwide, political, social, economic, and cultural contexts have greatly influenced the degree and nature of freedom national governments tend to grant their teachers in curriculum implementation. Their decisions about the selected combination of freedom from input and output control mechanisms show associations with patterns in how teachers behave in the classrooms considering their qualification characteristics, how students perform, and to what extent the systems manage to narrow down the gap in student achievement by socioeconomic background. Understanding the nuances of these associations can help identify the areas of risks or opportunities in the context of each type of autonomy and inform future policy interventions.

## REFERENCES

- Abe, T. O., & Adu, E. I. (2014). The effect of teachers' qualifications on students' performance in mathematics. *Sky Journal of Educational Research*, 2(1), 010-014.  
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f4a095a65b95206e37c02ffc0830eec3ddbbc7d0>.
- Alley, K. M. (2019). Fostering middle school students' autonomy to support motivation and engagement. *Middle School Journal*, 50(3), 5-14.  
<https://doi.org/10.1080/00940771.2019.1603801>.
- Amin, M. R. (2019). Developing reading skills through effective reading approaches. *International Journal of Social Science and Humanities*, 4(1), 35-40. <https://doi.org/10.5281/zenodo.2557919>.
- Anderson, L. W. (1987). The decline of teacher autonomy: Tears or cheers?. *International Review of Education*, 33(3), 357-373.  
<https://doi.org/10.1007/bf00615308>.
- Andersson, U., Cuervo-Cazurra, A., & Nielsen, B. B. (2020). Explaining interaction effects within and across levels of analysis. *Research Methods in International Business*, 331-349. <http://doi:10.1057/jibs.2014.50>.
- Ashley, V. (2012). Philosophical Models of Personal Autonomy.  
<https://autonomy.essex.ac.uk/wp-content/uploads/2016/11/Essex-Autonomy-Project-Philosophical-Models-of-Autonomy-October-2012.pdf>.
- Assor, A., Kaplan, H., Kanat-Maymon, Y., & Roth, G. (2005). Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: The role of anger and anxiety. *Learning and instruction*, 15(5), 397-413.  
<https://doi.org/10.1016/j.learninstruc.2005.07.008>.
- Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap?. *Learning and Instruction*, 63, 101211.  
<https://doi.org/10.1016/j.learninstruc.2019.05.008>.
- Avanzi, L., Miglioretti, M., Velasco, V., Balducci, C., Vecchio, L., Fraccaroli, F., & Skaalvik, E. M. (2013). Cross-validation of the Norwegian teacher's self-efficacy scale (NTSES). *Teaching and Teacher Education*, 31, 69-78.  
<https://doi.org/10.1016/j.tate.2013.01.002>.



- Ayubayeva, N. (2018). *Teacher collaboration for professional learning: Case studies of three schools in Kazakhstan* [Doctoral dissertation, University of Cambridge]. University of Cambridge Repository.
- Bakken, A. S. (2019). Questions of autonomy in English teachers' discursive practices. *Educational Research*, 61(1), 105-122.  
<https://doi.org/10.1080/00131881.2018.1561202>.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variables *distinction in social psychological research: Conceptual, strategic and statistical consideration*. *Journal of Personality and Social Psychology*, 51, 1173-1182.  
<https://doi.org/10.1037/0022-3514.51.6.1173>.
- Basu, R. (2004). The rationalization of neoliberalism in Ontario's public education system, 1995–2000. *Geoforum*, 35(5), 621-634.  
<https://doi.org/10.1016/j.geoforum.2004.03.003>.
- Benson, P. (2011). What's new in autonomy. *The Language Teacher*, 35(4), 15-18.  
<https://doi.org/10.37546/jaltlt35.4-4>.
- Benson, P. (2016). Teacher autonomy and teacher agency. In G. Barkhuizen (Ed.), *Reflections on language teacher identity research* (18-23). Routledge.
- Berliner, D. C., & Glass, G. V. (Eds.). (2014). *50 myths and lies that threaten America's public schools: The real crisis in education*. Teachers College Press.
- Berry, B., Daughtrey, A., & Wieder, A. (2010). *Teacher leadership: Leading the way to effective teaching and learning*. Center for Teaching Quality.
- Biesta, G. (2015a). What is education for? On good education, teacher judgement, and educational professionalism. *European Journal of Education*, 50(1), 75-87.  
<https://doi.org/10.1111/ejed.12109>.
- Biesta, G. (2015b). *How does a competent teacher become a good teacher? On judgement, wisdom and virtuosity in teaching and teacher education*. John Wiley & Sons. <https://doi.org/10.1002/9781118977859.ch1>.
- Biesta, G., & Tedder, M. (2006). How is agency possible? Towards an ecological understanding of agency-as-achievement. Learning lives: Learning, identity, and agency in the life course.  
[https://www.researchgate.net/profile/Michael\\_Tedder/publication/228644383\\_How\\_is\\_agency\\_possible\\_Towards\\_an\\_ecological\\_understanding\\_of\\_agency-as-achievement/links/00b4952cadd9bd2b6a000000.pdf](https://www.researchgate.net/profile/Michael_Tedder/publication/228644383_How_is_agency_possible_Towards_an_ecological_understanding_of_agency-as-achievement/links/00b4952cadd9bd2b6a000000.pdf).

- Blank, & de las Alas, N. (2010). *Effects of Teacher Professional Development on Gains in Student Achievement : How Meta Analysis Provides Scientific Evidence Useful to Education Leaders*. Distributed by ERIC Clearinghouse.
- Blatchford, P., Bassett, P., Goldstein, H., & Martin, C. (2003). Are class size differences related to pupils' educational progress and classroom processes? findings from the institute of education class size study of children aged 5–7 years. *British Educational Research Journal*, 29(5), 709-730. <https://doi.org/10.1080/0141192032000133668>.
- Boote, D. N. (2006). Teachers' professional discretion and the curricula. *Teachers and Teaching: Theory and practice*, 12(4), 461-478. <https://doi.org/10.1080/13450600600644319>.
- Boyadzhieva, E. (2016). Learner-centered teaching and learner autonomy. *Procedia-Social and Behavioral Sciences*, 232, 35-40. <https://doi.org/10.1016/j.sbspro.2016.10.008>.
- Bucelli, I. (2017). The Role of 'Autonomy' in Teaching Expertise. *Journal of Philosophy of Education*, 51(3), 588–604. <https://doi.org/10.1111/1467-9752.12252>.
- Bunker, V. J. (2008). *Professional learning communities, teacher collaboration, and student achievement in an era of standards based reform*. [Doctoral Dissertation, Lewis and Clark College]. ProQuest Dissertations and Theses Global.
- Burden, P. R. (1981). Teachers' perceptions of their personal and professional development. <https://files.eric.ed.gov/fulltext/ED210258.pdf>.
- Carlgren, I., & Klette, K. (2008). Reconstructions of Nordic Teachers: Reform policies and teachers' work during the 1990s. *Scandinavian Journal of Educational Research*, 52(2), 117-133. <https://doi.org/10.1080/00313830801915754> .
- Chaaban, Y., Arar, K., Sawalhi, R., Alhouti, I., & Zohri, A. (2021). Exploring teachers' professional agency within shifting educational contexts: A comparative study of Lebanon, Qatar, Kuwait, and Morocco. *Teaching and Teacher Education*, 106, 103451. <https://doi.org/10.1016/j.tate.2021.103451>.
- Charters Jr, W. W. (1976). Sense of Teacher Work Autonomy: Measurement & Findings.
- Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review*, 84(3), 517-544. <https://doi.org/10.1177/0003122419847165>.

- Cimentada, J. (2020). A new perspective on the international achievement gap: is academic autonomy good for everyone?.  
file:///Users/kchach/Downloads/achievement\_gap%20(2).pdf.
- Codd, J. (2005). Teachers as 'managed professionals' in the global education industry: The New Zealand experience. *Educational Review*, 57(2), 193-206.  
<https://doi.org/10.1080/0013191042000308369>.
- Cribb, A., & Gewirtz, S. (2007). Unpacking autonomy and control in education: Some conceptual and normative groundwork for a comparative analysis. *European Educational Research Journal*, 6(3), 203-213.  
<https://doi.org/10.2304/eej.2007.6.3.203>.
- Crocco, M. S., & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability urban educators speak out. *Urban Education*, 42(6), 512-535. <https://doi.org/10.1177/0042085907304964>.
- Davies, B., & Bansel, P. (2007). Neoliberalism and education. *International journal of qualitative studies in education*, 20(3), 247-259.  
<https://doi.org/10.1080/09518390701281751>.
- Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, 29(1), 1-19.  
<https://doi.org/10.1007/s10869-013-9308-7>.
- Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International journal of educational research*, 37(8), 677-692.  
[https://doi.org/10.1016/s0883-0355\(03\)00065-x](https://doi.org/10.1016/s0883-0355(03)00065-x).
- Day, C. (2020). How teachers' individual autonomy may hinder students' academic progress and attainment: Professionalism in practice. *British Educational Research Journal*, 46(1), 247-264.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268.  
[https://doi.org/10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01).
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, 49(3), 182.  
<https://doi.org/10.1037/a0012801>.
- Den Brok, P., Bergen, T., Stahl, R. J., & Brekelmans, M. (2004). Students' perceptions of teacher control behaviors. *Learning and Instruction*, 14(4), 425-443.  
doi:10.1016/j.learninstruc.2004.01.004.

- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage Publications.
- Dewey, J. (1903). Democracy in education. *The elementary school teacher*, 4(4), 193-204.
- Duke, N. K., Ward, A. E., & Pearson, P. D. (2021). The science of reading comprehension instruction. *The Reading Teacher*, 74(6), 663-672. <http://doi:10.1002/trtr.1993>.
- Dwiningtiyas, G. N., Sofyan, D., & Puspita, H. (2020). Teachers' strategies in teaching reading comprehension. *Journal of Applied Linguistic and Literacy*, 4(2), 66-77. <https://doi.org/10.25157/jall.v4i2.3682>.
- Dworkin, G. (2015). The nature of autonomy. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28479. <https://doi.org/10.3402/nstep.v1.28479>.
- Egodawatte, G., McDougall, D., & Stoilescu, D. (2011). The effects of teacher collaboration in Grade 9 Applied Mathematics. *Educational Research for Policy and Practice*, 10(3), 189-209. <http://doi.org/10.1007/s10671-011-9104-y>.
- Erss, M. (2018). 'Complete freedom to choose within limits'—teachers' views of curricular autonomy, agency and control in Estonia, Finland and Germany. *The Curriculum Journal*, 29(2), 238-256. <https://doi.org/10.1080/09585176.2018.1445514>.
- Erss, M., Kalmus, V., & Autio, T. H. (2016). 'Walking a fine line': teachers' perception of curricular autonomy in Estonia, Finland and Germany. *Journal of Curriculum Studies*, 48(5), 589-609. <https://doi.org/10.1080/00220272.2016.1167960>.
- Erss, M., & Kalmus, V. (2018). Discourses of teacher autonomy and the role of teachers in Estonian, Finnish and Bavarian teachers' newspapers in 1991–2010. *Teaching and Teacher Education*, 76, 95-105. <https://doi.org/10.1016/j.tate.2018.08.003>.
- Etelapelto, A., Vahasantanen, K., Hokka, P., & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review*, 10, 45-65. <https://doi.org/10.1016/j.edurev.2013.05.001>.
- Etelapelto, A., Vahasantanen, K., & Hokka, P. (2015). How do novice teachers in Finland perceive their professional agency?. *Teachers and Teaching*, 21(6), 660–680. <https://doi.org/10.1080/13540602.2015.1044327>.

- Evans, M., & Boucher, A. R. (2015). Optimizing the power of choice: Supporting student autonomy to foster motivation and engagement in learning. *Mind, Brain, and Education*, 9(2), 87-91. <https://doi.org/10.1111/mbe.12073>.
- Forsberg, E., & Wermke, W. (2012). Knowledge sources and autonomy: German and Swedish teachers' continuing professional development of assessment knowledge. *Professional Development in Education*, 38(5), 741-758. <https://doi.org/10.1080/19415257.2012.694369>.
- Fox, S. (2012). Adolescents, graduated autonomy, and genetic testing. *Genetics Research International*, 2012. <https://doi.org/10.1155/2012/946032>.
- Foy, P. (2018). *PIRLS 2016: User Guide for the International Database*. International Association for the Evaluation of Educational Achievement (IEA).
- Freidson, E. (2001). *Professionalism, the third logic: On the practice of knowledge*. University of Chicago Press.
- Friedman, I. A. (1999). Teacher-perceived work autonomy: The concept and its measurement. *Educational and Psychological Measurement*, 59(1), 58-76. <https://doi.org/10.1177/00131649921969749>.
- Frostenson, M. (2015). Three forms of professional autonomy: de-professionalisation of teachers in a new light. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28464. <https://doi.org/10.3402/nstep.v1.28464>.
- Gajda, R., & Koliba, C. J. (2008). Evaluating and improving the quality of teacher collaboration: A field-tested framework for secondary school leaders. *Nassp Bulletin*, 92(2), 133-153. <https://doi.org/10.1177/0192636508320990>.
- Gobby, B., Keddie, A., & Blackmore, J. (2018). Professionalism and competing responsibilities: Moderating competitive performativity in school autonomy reform. *Journal of Educational Administration and History*, 50(3), 159-173. <https://doi.org/10.1080/00220620.2017.1399864>.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Distributed by ERIC Clearinghouse.
- Gorman, E. H., & Sandefur, R. L. (2011). "Golden age," quiescence, and revival: how the sociology of professions became the study of knowledge-based work. *Work and Occupations*, 38(3), 275-302. <https://doi.org/10.1177/0730888411417565>.
- Grant, A., Hann, T., Godwin, R., Shackelford, D., & Ames, T. (2020). A Framework for graduated teacher autonomy: Linking teacher proficiency with Autonomy. *The*

- Educational Forum*, 84(2), 100-113.  
Routledge. <https://doi.org/10.1080/00131725.2020.1700324>.
- Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research*, 76(2), 173-208. <https://doi.org/10.3102/00346543076002173>.
- Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., & Morrison, F. J. (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. *The Elementary School Journal*, 113(1), 3-24. <https://doi.org/10.1086/665816>.
- Gurganious, N. (2017). *The relationship between teacher autonomy and middle school students' achievement in science* [Doctoral dissertation, Walden University]. ProQuest Dissertations and Theses Global.
- Guskey, T. R., & Sparks, D. (2004). Linking professional development to improvements in student learning. *Research linking teacher preparation and student performance*, 12(11). <https://files.eric.ed.gov/fulltext/ED464112.pdf>.
- Gwaltney, K. D. (2012). *Teacher autonomy in the United States: establishing a standard definition, validation of a nationally representative construct and an investigation of policy affected teacher groups* [Doctoral dissertation, University of Missouri-Columbia]. ProQuest Dissertations and Theses Global.
- Haapaniemi, J., Venäläinen, S., Malin, A., & Palojoki, P. (2021). Teacher autonomy and collaboration as part of integrative teaching—Reflections on the curriculum approach in Finland. *Journal of Curriculum Studies*, 53(4), 546-562. <https://doi.org/10.1080/00220272.2020.1759145>.
- Hamilton, L. S., Stecher, B. M., & Klein, S. P. (Eds.). (2002). *Making sense of test-based accountability in education*. RAND.
- Hanushek, E. A. (2019). Testing, accountability, and the American economy. *The ANNALS of the American Academy of Political and Social Science*, 683(1), 110-128. <https://doi.org/10.1177/0002716219841299>.
- Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, 104, 212-232. <https://doi.org/10.1016/j.jdeveco.2012.08.002>.
- Hanushek, E. A., Kain, J., O'Brien, D., & Rivkin, S. G. (2005). *The market for teacher quality*. National Bureau of Economic Research.

- Hargreaves, A. (2000). Four ages of professionalism and professional learning. *Teachers and Teaching*, 6(2), 151-182. <https://doi.org/10.1080/713698714>.
- Hargreaves, A., & Shirley, D. (2009). *The fourth way. The inspiring future for educational change*. Thousand Oaks.
- Hargreaves, A., & Fullan, M. (2015). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Hargreaves, E., Berry, R., Lai, Y. C., Leung, P., Scott, D., & Stobart, G. (2013). Teachers' experiences of autonomy in continuing professional development: Teacher learning communities in London and Hong Kong. *Teacher Development*, 17(1), 19-34. <https://doi.org/10.1080/13664530.2012.748686>.
- Harris, D. N., & Sass, T. R. (2007). Teacher Training, Teacher Quality, and Student Achievement. Working Paper 3. *National Center for Analysis of Longitudinal Data in Education Research*. <http://doi.org/10.1007/s10780-010-9106-3>.
- Hattie, J. (2009). *Visible Learning*. Routledge.
- Heacox, D. (2018). *Making differentiation a habit: How to ensure success in academically diverse classrooms*. Free Spirit Publishing.
- Helgoy, I., & Homme, A. (2007). Towards a new professionalism in school? A comparative study of teacher autonomy in Norway and Sweden. *European Educational Research Journal*, 6(3), 232-249. <https://doi.org/10.2304/eeerj.2007.6.3.232>.
- Hermansen, H. (2017). Knowledge relations and epistemic infrastructures as mediators of teachers' collective autonomy. *Teaching and Teacher Education*, 65, 1-9. <https://doi.org/10.1016/j.tate.2017.03.003>.
- Hong, W. P., & Youngs, P. (2016). Why are teachers afraid of curricular autonomy? Contradictory effects of the new national curriculum in South Korea. *Asia Pacific Journal of Education*, 36(sup1), 20-33. <https://doi.org/10.1080/02188791.2014.959471>.
- Hood, N. (2020). Learning from lockdown: What the experiences of teachers, students and parents can tell us about what happened and where to next for New Zealand's school system. *The Education Hub*. <https://theeducationhub.org.nz/wp-content/uploads/2020/08/Learning-from-lockdown.pdf>.
- Hopmann, S. T. (2003). On the evaluation of curriculum reforms. *Journal of Curriculum Studies*, 35(4), 459-478. <https://doi.org/10.1080/00220270305520>.

- Huberman, M. (1989). *The professional life cycle of teachers*. Teachers College Record.
- Huberman M. (1992). Teacher development and instructional mastery. In Hargreaves A., Fullan M. G. (Eds.), *Understanding teacher development* (pp. 122-142). New York, NY: Teachers College Press.
- Hudson, C. (2007). Governing the governance of education: The state strikes back?. *European Educational Research Journal*, 6(3), 266-282. <https://doi.org/10.2304/eej.2007.6.3.266> .
- Hulpia, H., Devos, G., & Van Keer, H. (2009). The influence of distributed leadership on teachers' organizational commitment: A multilevel approach. *The Journal of Educational Research*, 103(1), 40-52. <https://doi.org/10.1080/00220670903231201>.
- Hursh, D. (2005). Neo-liberalism, markets and accountability: Transforming education and undermining democracy in the United States and England. *Policy Futures in Education*, 3(1), 3-15. <https://doi.org/10.2304/pfie.2005.3.1.6>.
- Hyslop-Margison, E. J., & Sears, A. M. (2010). Enhancing teacher performance: The role of professional autonomy. *Interchange*, 41(1), 1-15. <https://doi.org/10.1007/s10780-010-9106-3> .
- Ingersoll, R. M. (1996). Teachers' decision-making power and school conflict. *Sociology of education*, 159-176. <https://doi.org/10.2307/2112804>.
- Ingersoll, R. M. (1997). *Teacher professionalization and teacher commitment: a multilevel analysis*, SASS. Department of Education Office of Educational.
- Johnson, B. (2003). Teacher collaboration: Good for some, not so good for others. *Educational Studies*, 29(4), 337-350. <https://doi.org/10.1080/0305569032000159651>.
- Johnson, J. A. (2016). TAKING THE “SQUEEZE” OFF OF SOCIAL STUDIES: A Phenomenological Study of Teacher “Curricular” Autonomy and its Effects on Social Studies Instructional Time. [Doctoral dissertation, University of Missouri-St. Louis]. UMSL Library.
- Juntunen, M. L. (2017). National assessment meets teacher autonomy: national assessment of learning outcomes in music in Finnish basic education. *Music Education Research*, 19(1), 1-16. <https://doi.org/10.1080/14613808.2015.1077799>.



- Kauffman, D., Johnson, S. M., Kardos, S. M., Liu, E., & Peske, H. G. (2002). Lost at sea: Without a curriculum, navigating instruction can be tough - especially for new teachers. *American Educator*, 26(2), 6. <https://doi.org/10.1111/1467-9620.00163>.
- Ke, Z., Yin, H., & Huang, S. (2019). Teacher participation in school-based professional development in China: does it matter for teacher efficacy and teaching strategies?. *Teachers and Teaching*, 25(7), 821-836. <https://doi.org/10.1080/13540602.2019.1662777>.
- Kelly, C. E. (2019). Teacher Perceptions of an International Baccalaureate Diploma Program in a Mid-western Inner-ring Suburban High School During the First Seven Years. [Doctoral Dissertation, Cleveland State University]. ProQuest Dissertations and Theses Global.
- Kennedy, M. M. (1999). Approximations to indicators of student outcomes. *Educational Evaluation and Policy Analysis*, 21(4), 345-363. <https://doi.org/10.3102/01623737021004345>.
- Khachatryan, S., Petrosyan, S., & Terzyan, G. (2013). Assessment of teacher professional development and educational content in the context of general education reforms in Armenia. <https://www.osf.am/wp-content/uploads/2014/03/FinalENGAssessmentPDEC.pdf>.
- Klassen R. M., Chiu M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, 102, 741-756. <https://doi.org/10.1037/a0019237>.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed). Guilford.
- Küçükoğlu, H. (2013). Improving reading skills through effective reading strategies. *Procedia-Social and Behavioral Sciences*, 70, 709-714. <https://doi.org/10.1016/j.sbspro.2013.01.113>.
- LaCoe, C. S. (2006). *Decomposing teacher autonomy: A study investigating types of teacher autonomy and how current public school climate affects teacher autonomy*. [Doctoral Dissertation, University of Pennsylvania]. ProQuest Dissertations and Theses Global.
- LaCoe, C. S. (2008). *Teacher autonomy: a multifaceted approach for the new millennium*. Cambria Press.

- Ladd, H. F. (2008). Teacher effects: What do we know. In G. Duncan, & J. Spillane (Eds.), *Teacher quality: Broadening and deepening the debate*, (pp. 3-26). Northwestern University.
- Lambeth, D. (2012). Effective practices and resources for support of beginning teachers. *Academic Leadership: The Online Journal*, 10(1), 2. <https://doi.org/10.58809/ALJ20120101/AQLU7946>.
- LaRoche, S., Joncas, M., & Foy, P. (2017). Chapter 3: Sample Design in PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016* (pp. 3.1-3.34). International Association for the Evaluation of Educational Achievement (IEA). <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-3.html>.
- Lawson, T. (2004). Teacher autonomy: Power or control?. *Education*, 3-13, 32(3), 3-18. <https://doi.org/10.1080/03004270485200261>.
- Lay, Y. F., & Chandrasegaran, A. L. (2016). Availability of School Resources and TIMSS Grade 8 Students' Science Achievement: A Comparative Study between Malaysia and Singapore. *International Journal of Environmental and Science Education*, 11(9), 3065-3080. <http://doi.org/10.12973/ijese.2016.907a>.
- Leana, C. R., & Pil, F. K. (2006). Social capital and organizational performance: Evidence from urban public schools. *Organization Science*, 17(3), 353-366. <https://doi.org/10.1287/orsc.1060.0191>.
- Lennert da Silva, A. L., & Mølsted, C. E. (2020). Teacher autonomy and teacher agency: a comparative study in Brazilian and Norwegian lower secondary education. *The Curriculum Journal*, 31(1), 115-131. <https://doi.org/10.1002/curj.3>.
- Leroy, N., Bressoux, P., Sarrazin, P., & Trouilloud, D. (2007). Impact of teachers' implicit theories and perceived pressures on the establishment of an autonomy supportive climate. *European journal of psychology of education*, 22(4), 529-545. <https://doi.org/10.1007/bf03173470>.
- Levine, T. H., & Marcus, A. S. (2007). Closing the achievement gap through teacher collaboration: Facilitating multiple trajectories of teacher learning. *Journal of Advanced Academics*, 19(1), 116-138. <https://doi.org/10.4219/jaa-2007-707>.
- Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509-536.
- Littlewood, W. (1996). "Autonomy": An anatomy and a framework. *System*, 24(4), 427-435. [https://doi.org/10.1016/s0346-251x\(96\)00039-5](https://doi.org/10.1016/s0346-251x(96)00039-5).

- Lundström, U. (2015). Teacher autonomy in the era of New Public Management. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28144. <https://doi.org/10.3402/nstep.v1.28144>.
- Luthans, F. (1992). *Organizational behavior*. McGraw-Hill.
- MacDonald, C. (2002). Relational professional autonomy. *Cambridge Quarterly of Healthcare Ethics*, 11(3), 282-289. <https://doi.org/10.1017/s0963180102113090>.
- Manteaw, B. O. (2008). When businesses go to school: Neoliberalism and education for sustainable development. *Journal of Education for Sustainable Development*, 2(2), 119-126. <https://doi.org/10.1177/097340820800200209>.
- Marks, H. M., & Louis, K. S. (1997). Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educational Evaluation and Policy Analysis*, 19(3), 245-275. <https://doi.org/10.3102/01623737019003245>.
- Martin, M, Mullis, I. V.S., Hooper, M., Yin, L., Foy, P., Fishbein, B. & Liu, J. (2017). Chapter 14: Creating and Interpreting the PIRLS 2016 Context Questionnaire Scales. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016* (pp. 14.1-14.106). International Association for the Evaluation of Educational Achievement (IEA). <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-3.html>.
- Mausethagen, S. (2013). Accountable for what and to whom? Changing representations and new legitimation discourses among teachers under increased external control. *Journal of Educational Change*, 14(4), 423-444. <https://doi.org/10.1007/s10833-013-9212-y>.
- Mausethagen, S., & Molstad, C. E. (2015). Shifts in curriculum control: contesting ideas of teacher autonomy. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28520. <https://doi.org/10.3402/nstep.v1.28520>.
- Mayer, A. P., Donaldson, M. L., LeChasseur, K., Welton, A. D., & Cobb, C. D. (2013). Negotiating site-based management and expanded teacher decision making: A case study of six urban schools. *Educational Administration Quarterly*, 49(5), 695-731. <https://doi.org/10.1177/0013161x13492793>.
- McGrath, I. (2006). Teachers' and learners' images for coursebooks. *ELT Journal*, 60(2), 171-180. <https://doi.org/10.1093/elt/cci104>.

- McLaughlin, M. W., Pfeifer, R. S., Swanson-Owens, D., & Yee, S. (1986). Why teachers won't teach. *The Phi Delta Kappan*, 67(6), 420-426.
- Meirink, J. A. (2007). *Individual teacher learning in a context of collaboration in teams*. Leiden University.
- Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, 3(1), 1-11. [http://doi.org/10.47263/jasem.3\(1\)01](http://doi.org/10.47263/jasem.3(1)01).
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. Sage.
- Mills, D. L. (2016). *A quantitative causal-comparative study of teacher autonomy and the effect on student achievement* [Doctoral dissertation, University of Phoenix]. ProQuest Dissertations and Theses Global.
- Miningou, É. W., Pierre-Louis, M., & Bernard, J. M. (2022). Improving learning outcomes in francophone Africa: More resources or improved efficiency?. *African Development Review*, 34(1), 127-141. <http://doi.org/10.1111/1467-8268.12615>.
- Molander, A., Terum, L. I., & Smeby, J. C. (Eds.). (2008). *Profesjonsstudier*.
- Moomaw, W. E. (2005). *Teacher-perceived autonomy: A construct validation of the teacher autonomy scale*. [Doctoral dissertation, University of West Florida]. UWF Digital Library.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. International Association for the Evaluation of Educational Achievement (IEA). <http://timssandpirls.bc.edu/pirls2016/international-results/>.
- Mullis I. V., Martin, O.M., & Sainsbury, M. (2015). Chapter 1: PIRLS 2016 Reading Framework. In I.V. Mullis, M.O. Martin, & M. Sainsbury (Eds.). *PIRLS 2016 reading framework* (pp. 11-29). International Association for the Evaluation of Educational Achievement (IEA). [https://timss.bc.edu/pirls2016/downloads/P16\\_FW\\_Chap1.pdf](https://timss.bc.edu/pirls2016/downloads/P16_FW_Chap1.pdf).
- Murphy, J., Hallinger, P., & Lotto, L. S. (1986). Inequitable Allocations of Alterable Learning Variables1. *Journal of Teacher Education*, 37(6), 21-26. *Researcher*, 27(5), 345-354. <https://doi.org/10.1007/s40299-018-0391-4>.
- Mustafa, M., & Cullingford, C. (2008). Teacher autonomy and centralised control: The case of textbooks. *International Journal of Educational Development*, 28(1), 81-88. <https://doi.org/10.1016/j.ijedudev.2007.07.003>.

- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning?. *Education Policy Analysis Archives*, *14*, 1-175. <https://doi.org/10.14507/epaa.v14n1.2006>.
- O'Hara, D. P. (2006). *Teacher autonomy: Why do teachers want it, and how do principals determine who deserves it?* [Doctoral Dissertation, University of Pennsylvania]. UPENN Repository.
- Olivant, K. F. (2015). "I am not a format": Teachers' experiences with fostering creativity in the era of accountability. *Journal of Research in Childhood Education*, *29*(1), 115-129. <https://doi.org/10.1080/02568543.2014.978920>.
- Orakci, S., & Gelisli, Y. (2019). The Effect of the Application of Learning Activities Based on Learner Autonomy on the 6th Grade Students' English Achievements, Attitudes, and Learner Autonomy. *International Journal of Curriculum and Instruction*, *11*(1), 269-292.
- Ozturk, I. H. (2012). Teacher's role and autonomy in instructional planning: The case of secondary school history teachers with regard to the preparation and implementation of annual instructional plans. *Educational Sciences: Theory and Practice*, *12*(1), 295-299. <https://files.eric.ed.gov/fulltext/EJ978444.pdf>.
- Paradis, A., Lutovac, S., & Kaasila, R. (2015). A Canadian teacher's perceived autonomy and self-confidence in the midst of an educational reform. *Problems of Education in the 21st Century*, *66*(1), 42-53.
- Paradis, A., Lutovac, S., Jokikokko, K., & Kaasila, R. (2018). Canadian and Finnish upper-secondary school mathematics teachers' perceptions of autonomy. *Pedagogy, Culture & Society*, *26*(3), 381-396. <https://doi.org/10.1080/14681366.2017.1407957>.
- Paradis, A., Lutovac, S., Jokikokko, K., & Kaasila, R. (2019). Towards a relational understanding of teacher autonomy: The role of trust for Canadian and Finnish teachers. *Research in Comparative and International Education*, *14*(3), 394-411. <https://doi.org/10.1177/1745499919864252>.
- Paris, A., & Paris, S. G. (2007). Teaching narrative comprehension strategies to first graders. *Cognition and Instruction*, *25*(1), 1-44. <https://doi.org/10.1080/07370000709336701>.
- Parker, G. (2015). Teachers' autonomy. *Research in Education*, *93*(1), 19-33. <https://doi.org/10.7227/rie.0008>.

- Paulsrud, D., & Wermke, W. (2020). Decision-making in Context: Swedish and Finnish teachers' perceptions of autonomy. *Scandinavian Journal of Educational Research*, 64(5), 706-727. <https://doi.org/10.1080/00313831.2019.1596975>.
- Pearson, L. C., & Hall, B. W. (1993). Initial construct validation of the teaching autonomy scale. *The Journal of Educational Research*, 86(3), 172-178. <https://doi.org/10.1080/00220671.1993.9941155>.
- Pearson, L. C., & Moomaw, W. (2005). The relationship between teacher autonomy and stress, work satisfaction, empowerment, and professionalism. *Educational Research Quarterly*, 29(1), 38-54. <https://doi.org/10.3200/joer.100.1.44-51>.
- Pelletier, L. G., & Sharp, E. C. (2009). Administrative pressures and teachers' interpersonal behaviour in the classroom. *Theory and Research in Education*, 7(2), 174-183. <https://doi.org/10.1177/1477878509104322>.
- Perry, L. B., & McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record*, 112(4), 1137-1162. <https://doi.org/10.1177/0161468110112004>.
- Philippou, S., Kontovourki, S., & Theodorou, E. (2014). Can autonomy be imposed? Examining teacher (re) positioning during the ongoing curriculum change in Cyprus. *Journal of Curriculum Studies*, 46(5), 611-633. <https://doi.org/10.1080/00220272.2013.856033>.
- Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? A review of US research. *Journal of Professional Capital and Community*, 4(4), 286-308. <https://doi.org/10.1108/JPC-12-2018-0032>.
- Porter, A. C. (1989). External standards and good teaching: The pros and cons of telling teachers what to do. *Educational Evaluation and Policy Analysis*, 11(4), 343-356. <https://doi.org/10.3102/01623737011004343>.
- Prichard, C., & Moore, J. (2016). The balance of teacher autonomy and top-down coordination in ESOL programs. *TESOL Quarterly*, 50(1), 190-201. <https://doi.org/10.1002/tesq.278>.
- Priestley, M., Biesta, G., & Robinson, S. (2015a). *Teacher agency: An ecological approach*. Bloomsbury Publishing.
- Priestley, M., Biesta, G., Philippou, S. & Robinson, S. (2015b). The teacher and the curriculum: exploring teacher agency. In D. Wyse, L. Hayward, & J. Pandya

(Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment* (pp. 187-201). SAGE Publications Ltd.

- Ramos, R. C. (2006). Considerations on the role of teacher autonomy. *Colombian Applied Linguistics Journal*, 183-202. <https://doi.org/10.14483/22487085.10510>.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). Russell Sage Foundation.
- Ro, J. (2020). Curriculum, standards and professionalisation: The policy discourse on teacher professionalism in Singapore. *Teaching and Teacher Education*, 91, 103056. <https://doi.org/10.1016/j.tate.2020.103056>.
- Robertson, S. L. (2008). “Remaking the World” Neoliberalism and the Transformation of Education and Teachers’ Labor. In M. F. Compton, & L. Weiner (Eds.), *The global assault on teaching, teachers, and their unions stories for resistance* (pp. 11-27). Palgrave Macmillan.
- Ronnberg, L. (2007). A recent Swedish attempt to weaken State control and strengthen school autonomy: The experiment with local time schedules. *European Educational Research Journal*, 6(3), 214-231. <https://doi.org/10.2304/eej.2007.6.3.214>.
- Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, 107(9), 1999-2045. <https://doi.org/10.1111/j.1467-9620.2005.00583.x>.
- 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>.
- Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will?. *Journal of Personality*, 74(6), 1557-1586. <https://doi.org/10.1111/j.1467-6494.2006.00420.x>.
- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of educational change*, 11(1), 45-61. <https://doi.org/10.1007/s10833-008-9098-2>.
- Sahlberg, P. (2011). *Finnish lessons. What Can the World Learn from Educational Change in Finland?* Teachers College Press.

- Salokangas, M., & Ainscow, M. (2017). *Inside the autonomous school: Making sense of a global educational trend*. Routledge.
- Salokangas, M., & Wermke, W. (2020). Unpacking autonomy for empirical comparative investigation. *Oxford Review of Education*, 46(5), 563-581. <https://doi.org/10.1080/03054985.2020.1733514>.
- Salonen-Hakomaki, S. M., Soini, T., Pietarinen, J., & Pyhalto, K. (2016). The way ahead for Finnish comprehensive school? Examining state-level school administrators' theory of change. *Journal of curriculum studies*, 48(5), 671-691. <https://doi.org/10.1080/00220272.2016.1143530>.
- Sandsør, A. M. J., Zachrisson, H. D., Karoly, L. A., & Dearing, E. (2023). The Widening Achievement Gap Between Rich and Poor in a Nordic Country. *Educational Researcher*, 52(4), 195-205. <https://doi.org/10.3102/0013189X221142596>.
- Saw, G. K. (2016). Patterns and trends in achievement gaps in Malaysian Secondary Schools (1999–2011): gender, ethnicity, and socioeconomic status. *Educational Research for Policy and Practice*, 15, 41-54. <https://doi.org/10.1007/s10671-015-9175-2>.
- Scheurich, J. J., & Skrla, L. (Eds.). (2004). *Educational equity and accountability: Paradigms, policies and politics*. Psychology Press.
- Schumm, J. S., Moody, S. W., & Vaughn, S. (2000). Grouping for reading instruction: Does one size fit all?. *Journal of learning disabilities*, 33(5), 477-488. <https://doi.org/10.1177/00222194000330050>.
- Short, P. M., & Rinehart, J. S. (1992). School participant empowerment scale: Assessment of level of empowerment within the school environment. *Educational and Psychological Measurement*, 52(4), 951-960. <https://doi.org/10.1177/0013164492052004018>.
- Silova, I., & Kazimzade, E. (2010). The changing status of the teaching profession in post-Soviet Azerbaijan: Implications for teacher education. In K. Karras & C. C. Wolhuter (Eds.), *International handbook on teacher education worldwide* (pp. 55–72). Atrapos Editions.
- Silova, I. (2018). Comparing post-socialist transformations: Dead ends, new pathways and unexpected openings. In M. Chankseliani, & I. Silova (Eds.). *Comparing post-socialist transformations: purposes, policies, and practices in education* (pp. 193-199). Symposium Books.
- Silova, I., Millei, Z., Chachkhiani, K., Palandjian, G., Vitrukh, M. (2021). Post-Socialist Transformations in Comparative and International Education. In T. D. Jules, R.



- Shields, M. A.M. Thomas (Eds.). *The Bloomsbury handbook of theory in comparative and international education* (pp. 147-166). Bloomsbury Publishing.
- Simpson, A., Sang, G., Wood, J., Wang, Y., & Ye, B. (2018). A dialogue about teacher agency: Australian and Chinese perspectives. *Teaching and Teacher Education*, 75, 316-326. <https://doi.org/10.1016/j.tate.2018.07.002>.
- Sinclair, B. (2009). The teacher as learner: Developing autonomy in an interactive learning environment. *Maintaining control: Autonomy and language learning*, 175-198. <https://doi.org/10.5790/hongkong/9789622099234.003.0010>.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417-453. <https://doi.org/10.3102/00346543075003417>.
- Skaalvik, E. M., & Skaalvik, S. (2009). Does school context matter? Relations with teacher burnout and job satisfaction. *Teaching and Teacher Education*, 25(3), 518-524. doi:10.1016/j.tate.2008.12.006.
- Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relation with teacher engagement, job satisfaction and emotional exhaustion. Psychological exhaustion. *Psychological Reports*, 114(1) <https://doi.org/10.2466/14.02.pr0.114k14w0>.
- Slavit, D., Kennedy, A., Lean, Z., Nelson, T. H., & Deuel, A. (2011). Support for professional collaboration in middle school mathematics: A complex web. *Teacher Education Quarterly*, 38(3), 113-131.
- Smith, R., & Erdogan, S. (2008). Teacher-learner autonomy: Programme goals and student-teacher constructs. In L. Terry, & H. Reinders (Eds.), *Learner and teacher autonomy: Concepts, realities and responses* (pp. 83-102). John Benjamins.
- Smyth, J. (2001). *Critical politics of teachers' work: An Australian perspective*. Peter Lang Publishing.
- Solomon, D., Battistich, V., & Hom, A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. *The Journal of Experimental Education*, 64(4), 327-347. <https://doi.org/10.1080/00220973.1996.10806602>.
- Steiner-Khamsi, G., Harris-Van Keuren, C., Silova, I., & Chachkhiani, K. (2008). Global Monitoring Report 2009: UNESCO.

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.6649&rep=rep1&type=pdf>.

Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist, 39*(2), 97-110.  
[http://doi:10.1207/s15326985ep3902\\_2](http://doi:10.1207/s15326985ep3902_2).

Stephens, P., Egil tonnessen, F., & Kyriacou, C. (2004). Teacher training and teacher education in England and Norway: a comparative study of policy goals. *Comparative Education, 40*(1), 109-130.  
<https://doi:10.1080/0305006042000184908>.

Stienstra, K., Knigge, A., Maas, I., de Zeeuw, E. L., & Boomsma, D. I. (2022). Are classrooms equalizers or amplifiers of inequality? A genetically informative investigation of educational performance. *European Sociological Review*.  
<https://doi.org/10.1093/esr/jcac054>.

Strong, L. E., & Yoshida, R. K. (2014). Teachers' autonomy in today's educational climate: Current perceptions from an acceptable instrument. *Educational Studies, 50*(2), 123-145. <https://doi.org/10.1080/00131946.2014.880922>.

Tahirsylaj, A. (2019). Teacher autonomy and responsibility variation and association with student performance in Didaktik and curriculum traditions. *Journal of Curriculum Studies, 51*(2), 162-184.  
<https://doi.org/10.1080/00220272.2018.1535667>.

Takala, T., & Piattoeva, N. (2012). Changing conceptions of development assistance to education in the international discourse on post-Soviet countries. *International Journal of Educational Development, 32*(1), 3-10.  
<https://doi.org/10.1016/j.ijedudev.2010.10.003>.

Teleshaliyev, N. (2013). "Leave Me Alone—Simply Let Me Teach" An Exploration of Teacher Professionalism in Kyrgyzstan. *European Education, 45*(2), 51-74.  
<https://doi.org/10.2753/eue1056-4934450203>.

Tetzlaff, L., Hartmann, U., Dumont, H., & Brod, G. (2022). Assessing individualized instruction in the classroom: Comparing teacher, student, and observer perspectives. *Learning and Instruction, 82*, 101655.  
<https://doi.org/10.1016/j.learninstruc.2022.101655>.

The Economist Intelligence Unit (2017). Democracy index 2016. Revenge of the deplorables. <http://felipesahagun.es/wp-content/uploads/2017/01/Democracy-Index-2016.pdf>.

- The World Bank. (2022). *The World Bank atlas method – Detailed methodology*.  
<https://datahelpdesk.worldbank.org/knowledgebase/articles/378832-what-is-the-world-bank-atlas-method>.
- Tickle, L. (2000). *Teacher induction: The way ahead*. Open University Press.
- Tieso, C. (2005). The effects of grouping practices and curricular adjustments on achievement. *Journal for the Education of the Gifted*, 29(1), 60-89.  
<https://doi.org/10.1177/016235320502900104>.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice?. *Theory into practice*, 44(3), 262-269. [https://doi.org/10.1207/s15430421tip4403\\_11](https://doi.org/10.1207/s15430421tip4403_11)
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners*. ASCD.
- Vangrieken, K., Grosemans, I., Dochy, F., & Kyndt, E. (2017). Teacher autonomy and collaboration: A paradox? Conceptualising and measuring teachers' autonomy and collaborative attitude. *Teaching and Teacher Education*, 67, 302-315.  
<https://doi.org/10.1016/j.tate.2017.06.021>.
- Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, 23(3), 263.  
<https://doi.org/10.1037/a0032359>.
- Verger, A., Fontdevila, C., & Zancajo, A. (2016). *The privatization of education: A political economy of global education reform*. Teachers College Press.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychological Bulletin*, 140(4), 1174.  
<https://doi.org/10.1037/a0036620>.
- Watts-Taffe, S., Laster, B. P., Broach, L., Marinak, B., McDonald Connor, C., & Walker-Dalhouse, D. (2012). Differentiated instruction: Making informed teacher decisions. *The Reading Teacher*, 66(4), 303-314.  
<https://doi.org/10.1002/TRTR.01126>.
- Webb, P. T. (2002). Teacher power: The exercise of professional autonomy in an era of strict accountability. *Teacher Development*, 6(1), 47-62.  
<https://doi.org/10.1080/13664530200200156>.
- Wermke, W. (2013). *Development and autonomy: Conceptualising teachers' continuing professional development in different national contexts* [Doctoral dissertation, Stockholm University]. Digitala Vetenskapliga Arkivet.

- Wermke, W., & Hostfalt, G. (2014). Contextualizing teacher autonomy in time and space: A model for comparing various forms of governing the teaching profession. *Journal of Curriculum Studies*, 46(1), 58-80. <https://doi.org/10.1080/00220272.2013.812681>.
- Wermke, W., & Forsberg, E. (2017). The changing nature of autonomy: Transformations of the late Swedish teaching profession. *Scandinavian Journal of Educational Research*, 61(2), 155-168. <https://doi.org/10.1080/00313831.2015.1119727>.
- White, P. A. (1992). Teacher empowerment under “ideal” school-site autonomy. *Educational Evaluation and Policy Analysis*, 14(1), 69-82. <https://doi.org/10.2307/1164528>.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, 112(4), 1008-1037. <https://doi.org/10.1177/016146811011200408>.
- Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, 100, 61-78. <https://doi.org/10.1016/j.jpubeco.2013.01.006>.
- Woessmann, L. (2005). The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13(2), 143-169. <https://doi.org/10.1080/09645290500031165>.
- World Bank Data Team (2016). *New country classifications by income level: 2016-2017*. <https://blogs.worldbank.org/opendata/new-country-classifications-2016>.
- Wilches, J. U. (2007). Teacher autonomy: A critical review of the research and concept beyond applied linguistics. *Íkala, revista de lenguaje y cultura*, 12(18), 245-275.
- Wills, J. S., & Sandholtz, J. H. (2009). Constrained Professionalism: Dilemmas of Teaching in the Face of Test-Based Accountability. *Teachers College Record*, 111(4), 1065-1114. <https://doi.org/10.1177/01614681091110040>.
- Yang, M., & Lee, H. J. (2022). Do school resources reduce socioeconomic achievement gap? Evidence from PISA 2015. *International Journal of Educational Development*, 88, 102528. <https://doi.org/10.1016/j.ijedudev.2021.102528>.
- You, Y., & Morris, P. (2016). Imagining school autonomy in high-performing education systems: East Asia as a source of policy referencing in England. *Compare: A Journal of Comparative and International Education*, 46(6), 882-905. <https://doi.org/10.1080/03057925.2015.1080115>.

Zeng, Z. (2013). Pathways to pre-service teachers' professional development: Insights from teacher autonomy. In *International Academic Workshop on Social Science (IAW-SC-13)*. Atlantis Press. <https://doi.org/10.2991/iaw-sc.2013.194>.

Zuzovsky, R. (2009). Teachers' qualifications and their impact on student achievement: Findings from TIMSS 2003 data for Israel. *IERI Monograph Series. Issues and Methodologies in Large-Scale Assessments*, 2, 37-62.

APPENDIX A

SUMMARY OF MEASURES

Measure	Scale	Items/Variables incorporated
<b>Country-level</b>		
Teacher autonomy as freedom from control	<ul style="list-style-type: none"> <li>- Limited autonomy</li> <li>- Moderate input-driven autonomy</li> <li>- Moderate output-driven autonomy</li> <li>- Extended autonomy</li> </ul>	<p><b>What does the language/reading curriculum prescribe?</b> (<i>Yes; No</i>)</p> <ul style="list-style-type: none"> <li>- Goals and objectives (READ05B)</li> <li>- Instructional processes or methods (READ05C)</li> <li>- Materials (e.g., textbooks, instructional materials) (READ05D)</li> <li>- Assessment methods/activities (READ05E)</li> </ul> <p>Source: PIRLS 2016 curriculum questionnaire</p> <p><b>How is the language / reading curriculum implementation evaluated?</b> (<i>Yes; No</i>)</p> <ul style="list-style-type: none"> <li>- Visits by inspectors (READ06A)</li> <li>- Research programs (READ06B)</li> <li>- School self-evaluation (READ06C)</li> <li>- National or regional examinations (READ06D)</li> </ul> <p>Source: PIRSL 2016 curriculum questionnaire</p>
Country's economic development level	<ul style="list-style-type: none"> <li>- Low-income economy</li> <li>- Lower-middle-income economy</li> <li>- Upper-middle-income economy</li> <li>- High-income economy</li> </ul>	Source: World Bank's analytical classification of the world's economies
Country's democratic development level	<ul style="list-style-type: none"> <li>1 - Authoritarian regime</li> <li>2 - Hybrid regime</li> <li>3 - flawed democracy</li> <li>4 - Full democracy</li> </ul>	Source: The Economist Intelligence Unit's Democracy Index
Achievement gap in reading by quartiles	The difference between the average scores of students in the 75 <sup>th</sup> and 25 <sup>th</sup> percentiles	Source: PIRLS 2016 student questionnaire
<b>School-level</b>		
Share of economically disadvantaged students at school	<ul style="list-style-type: none"> <li>1 - 0-10 %</li> <li>2 - 11-25 %</li> <li>3 - 26-50 %</li> <li>4 - More than 50 %</li> </ul>	<p><b>Approximately what percentage of students in your school have the following backgrounds? (a) Come from economically disadvantaged homes.</b> (0-10 %; 11-25 %, 26-50 %, <i>More than 50 %</i>). ACBG03A.</p> <p>Source: PIRLS 2016 school questionnaire</p>
Instruction affected by reading resource shortages	Continuous (Range: 3.55 – 14.49)	<p><b>How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following?</b> (<i>Not at all, A little, Some, A lot</i>)</p> <p><u>A. General School Resources</u></p> <ul style="list-style-type: none"> <li>- Instructional materials (e.g., textbooks) (ACBG12AA)</li> <li>- Supplies (e.g., papers, pencils, materials) (ACBG12AB)</li> <li>- School buildings and grounds (ACBG12AC)</li> <li>- Heating/cooling and lighting systems (ACBG12AD)</li> <li>- Instructional space (e.g., classrooms) (ACBG12AE)</li> <li>- Technologically competent staff (ACBG12AF)</li> </ul>
	<ul style="list-style-type: none"> <li>1 - Not affected</li> <li>2 - Somewhat affected</li> <li>3 - Affected a lot</li> </ul>	

		<ul style="list-style-type: none"> <li>- Audio-visual resources for the delivery of instruction (e.g., interactive white boards, digital projectors) (ACBG12AG)</li> <li>- Computer technology for teaching and learning (e.g., computers or tablets for student use) (ACBG12AH)</li> </ul> <p><u>Resources for Reading Instruction</u></p> <ul style="list-style-type: none"> <li>- Teachers with a specialization in reading (ACBG12BA)</li> <li>- Computer software/ applications for reading instruction (ACBG12BB)</li> <li>- Library resources (books, ebooks, magazines, etc.) (ACBG12BC)</li> <li>- Instructional materials for reading (e.g., reading series, textbooks) (ACBG12BD)</li> </ul> <p>Source: PIRLS 2016 school questionnaire</p>
<b>Teacher-level</b>		
Teacher qualification: teacher education	<p>1 - Less than Bachelor's or equivalent</p> <p>2 - Bachelor's or equivalent</p> <p>3 - Bachelor's or equivalent</p>	<p><b>What is the highest level of formal education you have completed?</b> (<i>Circle one only</i>). ATBG04.</p> <ul style="list-style-type: none"> <li>- Did not complete upper secondary education-ISCED Level 3</li> <li>- Upper secondary education-ISCED Level 3</li> <li>- Post-secondary, non-tertiary education-ISCED Level 4</li> <li>- Short-cycle tertiary education-ISCED Level 5</li> <li>- Bachelor's or equivalent level-ISCED Level 6</li> <li>- Master's or equivalent level-ISCED Level 7</li> <li>- Doctor or equivalent level-ISCED Level 8</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher qualification: teaching experience	<p>Continuous (Range: 0 - 60)</p> <p>1 - Less than 5 years</p> <p>2 - At least 5 years but less than 10 years</p> <p>3 - At least 10 years but less than 20 years</p> <p>4 - 20 years or more.</p>	<p><b>By the end of this school year, how many years will you have been teaching altogether?</b> (<i>Please round to the nearest whole number</i>). ATBG01 / ATDG01.</p> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher qualification: formal professional development	<p>1 - None</p> <p>2 - Less than 6 hours</p> <p>3 - 6-15 hours</p> <p>4 - 16-35 hours</p> <p>5 - More than 35 hours</p>	<p><b>In the past two years, how many hours in total have you spent in formal professional development (e.g., workshops, seminars, lesson studies, etc.) that dealt directly with reading or teaching reading (e.g., reading theory, instructional methods)?</b> (<i>None, Less than 6 hours, 6-15 hours, 16-35 hours, More than 35 hours</i>). ATBG06.</p> <p>Source: PIRSL 2016 teacher questionnaire</p>
Teacher qualification: school-based professional development	<p>Continuous (Range: 4 - 16)</p> <p>1 - Never or almost never</p> <p>2 - Sometimes</p> <p>3 - Often</p> <p>4 - Very often</p>	<p><b>How often do you have the following types of interactions with other teachers?</b> (<i>Very often, Often, Sometimes, Never or Almost Never</i>).</p> <ul style="list-style-type: none"> <li>- Share what I have learned about my teaching experiences (ATBG09A)</li> <li>- Observe another classroom to learn more about teaching (ATBG09B)</li> </ul>



		<ul style="list-style-type: none"> <li>- Work together to improve how to teach a particular topic (ATBG09C)</li> <li>- Work with teachers from other grades to ensure continuity in learning (ATBG09E)</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher use of reading comprehension strategies	Continuous (Range: 9 - 36)	<p><b>How often do you ask the students to do the following things to help develop reading comprehension skills or strategies?</b> (<i>Every day or almost every day, Once or twice a week, Once or twice a month, Never or almost never</i>).</p> <ul style="list-style-type: none"> <li>- Locate information within the text (ATBR12A)</li> <li>- Identify the main ideas of what they have read (ATBR12B)</li> <li>- Explain or support their understanding of what they have read (ATBR12C)</li> <li>- Compare what they have read with experiences they have had (ATBR12D)</li> <li>- Compare what they have read with other things they have read (ATBR12E)</li> <li>- Make predictions about what will happen next in the text they are reading (ATBR12F)</li> <li>- Make generalizations and draw inferences based on what they have read (ATBR12G)</li> <li>- Describe the style or structure of the text they have read (ATBR12H)</li> <li>- Determine the author’s perspective or intention (ATBR12I)</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher use of reading content differentiation	Continuous (Range: 2- 8)	<p><b>How often do you do the following in teaching reading to this class?</b> (<i>Every or almost every lesson; About half the lessons; Some lessons; Never</i>).</p> <ul style="list-style-type: none"> <li>- Provide reading materials that match the students’ interests (ATBR11A)</li> <li>- Provide materials that are appropriate for the reading levels of individual students (ATBR11B)</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher use of autonomy-supportive strategies	Continuous (Range: 7- 28)	<p><b>When you have reading instruction and/or do reading activities, how often do you organize students in the following ways?</b> (<i>Always or almost always, Often, Sometimes, Never</i>). Source: Teacher questionnaire</p> <ul style="list-style-type: none"> <li>- Students work independently on an assigned plan or goal (ATBR8E)</li> </ul> <p><b>How often do you do the following in teaching reading to this class?</b> (<i>Every or almost every lesson, About half the lessons, Some lessons, Never</i>). Source: teacher questionnaire</p> <ul style="list-style-type: none"> <li>- Encourage students to develop their understandings of the text (ATBR11D)</li> </ul>

		<ul style="list-style-type: none"> <li>- Encourage student discussions of texts (ATBR11E)</li> <li>- Encourage students to challenge the opinion expressed in the text (ATBR11F)</li> <li>- Give students time to read books of their own choosing (ATBR11H)</li> <li>- Give individualized feedback to each student (ATBR11I)</li> </ul> <p><b>After students have read something, how often do you ask them to do the following?</b> (<i>Every or almost every lesson, About half the lessons, Some lessons, Never</i>). Source: Teacher questionnaire</p> <p>Talk with each other about what they have read (ATBR13C)</p> <p>Source: PIRLS 2016 teacher questionnaire</p>
Teacher use of individualized instruction	1 - Never or almost never; or sometimes 2 - Often, always or almost always.	<p><b>When you have reading instruction and/or do reading activities, how often do you organize students in the following</b> (<i>Always or almost always; Often; Sometimes; Never</i>).</p> <ul style="list-style-type: none"> <li>- I use individualized instruction for reading (ATBR08D)</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
Classroom size	Continuous (range: 1-89)	<p><b>How many students are in this class?</b> (<i>Write in the number</i>). ATBR01A.</p> <p>Source: PIRLS 2016 teacher questionnaire</p>
Instruction limited by student attributes	1 - Very little 2 – Somewhat 3 – A lot	<p><b>In your view, to what extent do the following limit how you teach this class?</b> (<i>A little, Some, A lot</i>).</p> <ul style="list-style-type: none"> <li>- Students lacking prerequisite knowledge or skills (ATBR05A)</li> <li>- Students suffering from lack of basic nutrition (ATBR05B)</li> <li>- Students suffering from not enough sleep (ATBR05C)</li> <li>- Students absent from class (ATBR05D)</li> <li>- Disruptive students (ATBR05E)</li> <li>- Uninterested students (ATBR05F)</li> <li>- Students with mental, emotional, or psychological impairment (ATBR05G)</li> </ul> <p>Source: PIRLS 2016 teacher questionnaire</p>
<b>Student-level</b>		
Student achievement in reading	Continuous (Range: 75.91 - 785.27)	<ul style="list-style-type: none"> <li>- Plausible Value 1: Overall reading (ASRREA01)</li> <li>- Plausible Value 2: Overall reading (ASRREA02)</li> <li>- Plausible Value 3: Overall reading (ASRREA03)</li> <li>- Plausible Value 4: Overall reading (ASRREA04)</li> <li>- Plausible Value 5: Overall reading (ASRREA05)</li> </ul>
Student Socio-economic background (Home resources for learning)	Continuous (Range: 4.03 - 14.8)	<p><b>Number of books in the home.</b> (<i>0–10, 11–25, 26–100, 101–200, more than 200</i>). ASBG04. Source: PIRLS 2016 student questionnaire</p>
	1- Few resources 2- Some resources	

	3- Many resources	<p><b>Number of home study supports.</b> ASDG05S (<i>None, Internet connection or own room, Both</i>). Source: PIRLS 2016 student questionnaire</p> <p><b>Number of children’s books in the home.</b> (0–10, 11–25, 26–100, 101–200, more than 200). ASBH14. Source: PIRLS 2016 parent questionnaire</p> <p><b>Highest level of education of either parents.</b> (<i>Finished some primary or lower secondary, or did not go to school, Finished lower secondary, finished upper secondary, finished pos-secondary education, finished university or higher</i>). ASDHEDUP. Source: PIRLS 2016 parent questionnaire</p> <p><b>Highest level of occupation of either parent.</b> (<i>Has never worked outside home for pay, general laborer or semi-professional; Clerical, Small-business owner, Professional</i>). ASDHOCCP. Source: PIRLS 2016 parent questionnaire</p>
Student gender	0 - Boy 1 - Girl	<p><b>Are you a girl or a boy?</b> (<i>Girl, Boy</i>). Source: PIRLS 2016 student questionnaire</p>