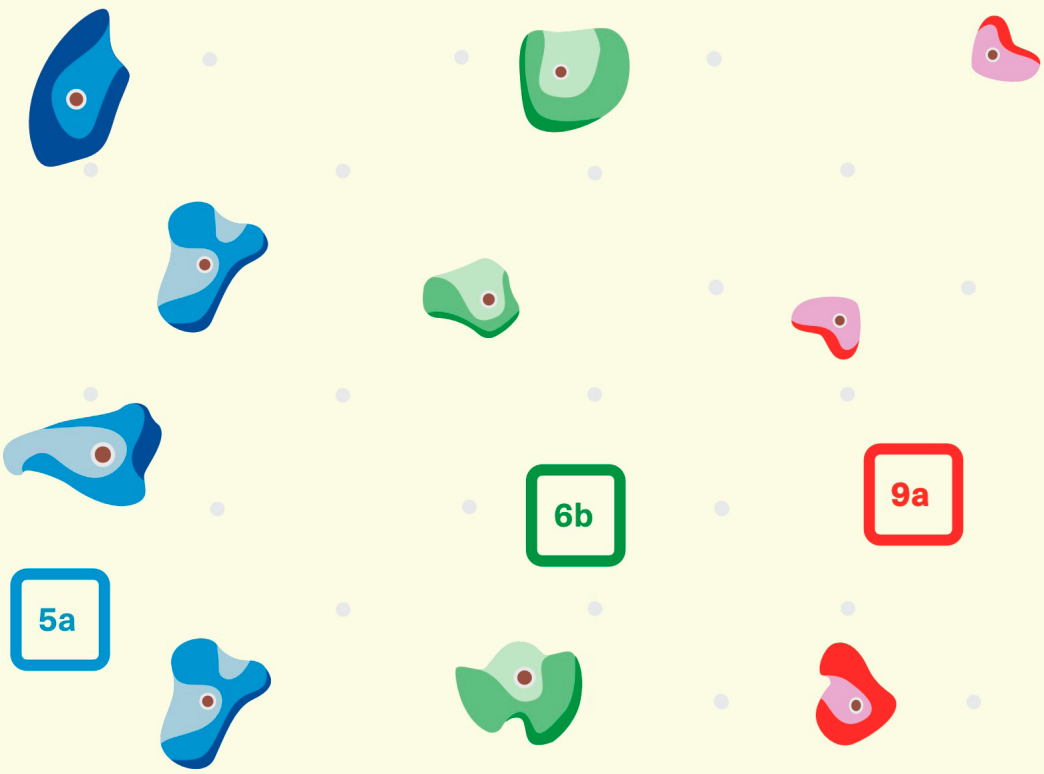


STUDIA PAEDAGOGICA
VOLUME 29 / NUMBER 3 / YEAR 2024

LEARNING OUTCOMES IN SCHOOLS: POLICIES AND PRACTICES

MASARYK UNIVERSITY PRESS

STUDIA PAEDAGOGICA
VOLUME 29 / NUMBER 3 / YEAR 2024

LEARNING OUTCOMES IN SCHOOLS: POLICIES AND PRACTICES

MARTIN SEDLÁČEK
PAULA KORŠŇÁKOVÁ
JANA STRAKOVÁ
EDITORS

2024

MASARYK
UNIVERSITY

Studia paedagogica cooperates with the Emerging Researchers' Group of the European Educational Research Association (EERA).

Citation: *Studia paedagogica* 29, 3, 2024

ISSN 2336-4521 (online)

CONTENTS

Editorial

Study

EDITORIAL

LEARNING OUTCOMES IN SCHOOLS: POLICIES AND PRACTICES

The concept of performance and achievement has long been a focal point of debate in the field of education. Over time, the focus has shifted toward learning outcomes, which now play a central role in shaping priorities and practices. Learning outcomes serve as a critical framework for redefining success in educational policy, emphasizing not only what individuals learn but also how they apply their knowledge in meaningful ways. In light of these evolving priorities, this special issue centers on learning outcomes, providing a platform for diverse perspectives and insights on their role in redefining educational success (Fuller, 2009; Prøitz, 2015).

Learning outcomes encompass multiple dimensions, including values, curriculum goals, and societal expectations (Hargreaves et al., 2001; Nygaard et al., 2009). They also bring to light critical questions about what constitutes "desirable" outcomes and how best to measure them (Mejía-Rodríguez & Kyriakides, 2022). For instance, critical thinking and adaptability are increasingly vital in navigating rapid technological advancements and global challenges, yet integrating these skills into curricula remains a persistent hurdle. Beyond academic knowledge, students today must acquire 21st-century skills such as problem-solving, cultural communication, and continuous learning, paired with ethical values, positive dispositions, and personal well-being.

International assessments, particularly PISA, have a significant impact on shaping educational policy (Hopmann, 2008; Schleicher, 2017). The focus of assessments largely determines what is prioritized in schools. Assessment results are used not only by learners and teachers to guide teaching and learning but also by researchers to examine systemic influences, interpret results within broader contexts—such as the outcomes and implications for different social groups—and address questions focused on the accurate measurement and interpretation of outcomes. Through the articles presented in this special issue, we aim to advance understanding of these critical questions and contribute to ongoing efforts to redefine success in education. For example, the first study, authored by Barbara Japelj Pavešić, Klaudija Šterman Ivančič, and Gašper Cankar, analyzes the broader context of educational outcomes. Titled *Achievement in the Light of Aspects of Student Well-being and Teacher Attitudes: The Case of Slovenia*, the study highlights the critical connection between educational outcomes and student well-being, drawing

on Slovenian data from international assessments. The findings reveal that student-teacher relationships significantly predict achievement, particularly in vocational and technical programs, emphasizing the importance of empowering teachers to create supportive and inclusive learning environments.

The second study, *If They Talk More During Lessons, Will They Achieve Better? Unlocking the Reciprocal Relationship Between Student Verbal Participation and Achievement*, was authored by Martin Sedláček, Klára Šeďová, Roman Švaříček, and Zuzana Šalamounová. The authors conducted an intervention study to examine the relationship between student verbal participation and achievement in sixth-grade language arts. While the intervention successfully increased and equalized student participation, no immediate impact on academic achievement was observed. The longitudinal nature of the data allowed the researchers to test the hypothesis that student talk time predicts achievement rather than the reverse. Their analysis confirmed that participation in classroom talk is indeed a significant predictor of achievement. The findings underscore the need for further research to investigate the long-term effects of increased verbal participation.

The third study, titled *From Policy to Practice: Monitoring SDG Targets with PIRLS 2021 Austrian Data*, by Surette van Staden and Sunet Grobler, investigate the alignment of Progress in International Reading Literacy Study (PIRLS) 2021 data with Sustainable Development Goal 4 (SDG4) targets, particularly focusing on quality education, equity, and lifelong learning opportunities. The findings highlight a decline in reading literacy in Austria and reveal notable gender disparities. The study further emphasizes the positive impact of early literacy activities and pre-primary education on reading outcomes, underscoring the critical role of early interventions. However, the analysis indicates that teacher qualifications have limited direct influence on reading achievement, pointing instead to the importance of teaching methods and classroom dynamics.

An example of a study focused on learning outcomes within a specific student population is the fourth study, titled *Analyzing Student Performance in Connection with Learning Disorders and Extracurricular Activities in Foreign Language Classes*. Authors Roland Hegedűs and Krisztina Sebestyén utilized data from the Hungarian National Assessment of Basic Competencies to examine the academic performance of students with atypical developmental pathways, including dyslexia, dysgraphia, and dyscalculia, compared to their peers without such disorders. The study explored how performance is influenced by family background and participation in different curricular programs. The findings reveal that students with learning disorders perform better in advanced foreign language classes than in standard curricula. However, advanced-level training in other subjects often leads to better learning outcomes, a trend frequently associated with variations in family background.

The last study, authored by Julie Grombířová and Jana Kratochvílová, adopts a completely different methodological approach. Titled *Innovation in the Final Assessment of Student Learning Outcomes (Report Cards): The Process of Change and its Impacts on School Practice*, the study employs a qualitative research design to examine the development and impact of implementing an innovative report cards in primary and lower secondary schools. Through interviews with 23 teachers and school leaders, the findings highlight the positive effects of the report card on teaching culture and assessment practices in primary schools, where it has been widely accepted by teachers. However, the study also reveals resistance among lower secondary teachers, pointing to the need for further development and dialogue to address these challenges.

The articles in this issue delve into the complexities surrounding the definition, implementation, and assessment of learning outcomes across varied educational contexts. We hope these contributions inspire readers to reflect on how learning outcomes can be harnessed to create equitable, effective, and meaningful educational opportunities. We invite you to explore these studies and consider how their insights might enrich your work and foster discussions within your professional communities.

Martin Sedláček, Paula Koršňáková and Jana Straková
editors of the special issue

References

Fuller, B. (2015). Policy and Place—Learning from Decentralized Reforms. In *Handbook of Education Policy Research*.
https://doi.org/10.4324/9780203880968.ch67

Hargreaves, A., Earl, L., Moore, S., & Manning, S. (2001). *Learning to change: Teaching beyond subjects and standards*. Jossey-Bass.

Mejía-Rodríguez, A. M., & Kyriakides, L. (2022). What matters for student learning outcomes? A systematic review of studies exploring system-level factors of educational effectiveness. *Review of Education*, *10*(3).
https://doi.org/10.1002/rev3.3374

Nygaard, C., Holtham, C., & Courtney, N. (Eds.) (2009). *Improving Students´ Learning Outcomes*. Copenhagen Business School Press.

Prøitz, T. S. (2015). Learning Outcomes as a Key Concept in Policy Documents throughout Policy Changes. *Scandinavian Journal of Educational Research*, *59*(3), 275–296.
https://doi.org/10.1080/00313831.2014.904418

Schleicher A. (2017). Seeing education through the prism of PISA. *European Journal of Education*. 52, 124–130.
https://doi.org/10.1111/ejed.12209

STUDIA PAEDAGOGICA

vol. 29, n. 3, 2024
https://doi.org/10.5817/SP2024-3-1

STUDY

ACHIEVEMENT IN THE LIGHT OF ASPECTS OF STUDENT WELL-BEING AND TEACHER ATTITUDES: THE CASE OF SLOVENIA

Barbara Japelj Pavešić[a], Klaudija Šterman Ivančič[a], Gašper Cankar[b]

[a] Educational Research Institute, Ljubljana
[b] National Examinations Centre, Ljubljana

ABSTRACT

Although many countries experienced negative trends in the 2022 cycle of PISA results, results for Slovenia showed an alarmingly negative trend in mathematics, reading, and science achievement on PISA literacy scales, and at the same time a decline in several aspects of student well-being. According to research evidence on the relatedness of knowledge and well-being, some studies have already confirmed student well-being as an important predictor of academic achievement. In the study, we focus on the differences in relationships between knowledge and aspects of student and teacher well-being in various secondary education programs. The survey results show that students and teachers in technical and vocational education programs report lower levels of some aspects of well-being, and that in explaining achievement in mathematics, reading and science, the most significant predictors across all education programs proved to be the quality of student-teacher relationships. The results suggest that by empowering teachers we can foster learning environments that support students from different backgrounds and as such enhance more positive student outcomes.

KEYWORDS

academic achievement; student well-being; teacher well-being; student outcomes

CORRESPONDING AUTHOR
Barbara Japelj Pavešić, Educational Research Institute, Gerbičeva 62, 1000 Ljubljana, Slovenia
e-mail: barbara.japelj@pei.si

Introduction

The recent findings from PISA 2022 for Slovenia reveal a concerning trajectory, characterized by a notable regression in mathematics, reading, and science proficiency, as documented by the OECD (2023a). The same cycle of PISA also reported several dimensions of student well-being (Šterman Ivančič & Mlekuž, 2023). This enables us to gain insight into the nature of educational outcomes and their association with other factors.

Researchers described the impact of teacher "quality" and "opportunity to learn" on student performance as modest, implying the existence of additional determinants (Carnoy et al., 2016). At the same time, the role of educators and the broader educational background warrants attention, as stated by Burris (2012), who attributes lower student performance in the U.S. to deficiencies in pedagogical training, learning environments, and the erosion of trust in teaching professionals. This is similar to Braun (2008), who states that successful educational systems are characterized by strong interrelations between student performance, school contexts, and classroom practices.

According to Baumert et al. (2009), large-scale assessments such as TIMSS and PISA measure the intricate processes of knowledge acquisition rather than solely assessing cognitive faculties. This means that those outcomes do not depend only on individual student characteristics outside the scope of teachers and schools, but rather include those as well. This means that students' outcomes do not depend only on individual student characteristics outside the scope of teachers and schools, but rather include characteristics of processes of knowledge acquisition as well.

That is why the decline in academic achievement in Slovenia necessitates a holistic outlook on many factors influencing student outcomes, including aspects of a teacher-student relationship, students' feelings of safety, belonging to school, and their self-oriented beliefs.

The role of **teacher-student relationships** in shaping both academic performance and student well-being has been extensively documented in educational research. Van Petegem et al. (2007) report a positive correlation between student well-being and interpersonal teacher behavior, emphasizing the profound influence that teacher-student dynamics have on student motivation and classroom climate. Similarly, Mikk et al. (2016) highlight the weak yet positive relationship between teacher-student relations and student motivation and academic performance. They stress the significance of fostering positive teacher-student interactions as a means to enhance overall school performance.

Perceived teacher competencies play a crucial role in shaping student emotions and well-being within educational contexts, particularly in subject-specific instruction. Gläser-Zikuda and Fuß (2008) report the impact of

perceived teacher competencies on student emotions in physics instruction. McGrath and Bergen (2015) report the nature of student-teacher relationships by identifying various factors such as student characteristics, periods of schooling at risk, and previous attachment relationships that influence the dynamics of these relationships.

The empirical evidence suggests that the quality of teacher-student relationship can serve as a predictor of academic outcomes and socio-emotional well-being among students. Hughes (2011) observes the longitudinal effects of teacher-student relationships on academic adjustment, revealing that high-quality teacher-student relationships predict positive academic self-images, behavioral engagement, and achievement in subsequent years. Importantly, student reports of teacher-student relationships predict feelings of school belonging and math achievement. By fostering positive teacher-student relationships we could promote both academic success and socio-emotional well-being of students.

Feeling safe within the school environment is another critical factor that influences both academic performance and well-being. School climate, staff actions, and interpersonal relationships all contribute to creating a sense of safety for students and faculty alike (Bosworth et al., 2011). Positive student-teacher relations, consistent disciplinary practices, and a supportive school community are essential components in fostering feelings of safety (Williams et al., 2018). Feeling safe is also important in the context of bullying, a closely connected concept that contributes to well-being (Volk et al., 2017). Importantly, perceptions of safety are not uniform across student groups, with variations observed based on gender, ethnicity, and socioeconomic status (Siann et al., 1994). Because of that, efforts to enhance school safety must consider the diverse needs and experiences of all students.

The concept of **school belonging** also emerges as a predictor of academic achievement and socio-emotional outcomes. Students who feel connected to their school community exhibit higher levels of motivation, engagement, and academic performance (Korpershoek et al., 2020). A sense of belonging is particularly crucial for students at risk of dropout, as it serves as a protective factor against disengagement and academic failure (Goodenow & Grady, 1993). Additionally, school belonging is positively associated with final-semester academic grades (Roeser et al., 1996). However, there are differences in the prevalence of school belonging across students with different socio-economic backgrounds, with higher levels observed in schools serving middle and high socio-economic status populations (Sari, 2012).

1 Contemporary studies

In Slovenia, international independent studies such as PISA, TALIS, and TIMSS consistently identify problems with the quality of relations in schools and attitudes toward learning. According to the latest PISA report, student achievement is declining, with a continuing downward trend in many students' attitudes toward learning, school, and relations. To improve both, focused research on the relationship between attitudes and knowledge is needed within the existing national initiative to provide a safe and supportive learning environment for all students.

Upper secondary education in Slovenia is tracked, and students are enrolled in programs of their own choice. Tracking leads to different future career decisions and therefore some differences across levels and types of programs or secondary schools are to be expected. The practice of tracking in several educational systems has gained attention due to its perceived influence on student outcomes and social inequalities. Research by Strello et al. (2021) indicates that while tracking may exacerbate social achievement gaps and dispersion inequalities, it does not necessarily lead to improved performance levels in assessments such as PISA or TIMSS. This suggests that early tracking may contribute to widening disparities in student achievement without yielding commensurate gains in academic proficiency across different tracks.

In their overview, Strello et al. (2021) note that the implications of early tracking extend to the differential experiences and opportunities afforded to students in different tracks. While some argue that tracking allows for tailored instruction and targeted support, others caution against the perpetuation of inequalities and limited mobility between tracks. The delineation of students into distinct tracks based on perceived ability or aptitude may inadvertently reinforce socio-economic disparities and limit upward mobility for students from disadvantaged backgrounds.

The effects of early tracking extend beyond academic performance and include learning inequalities, particularly in reading literacy. Contini and Cugnata (2020) highlight the role of early tracking in increasing overall inequalities and amplifying differences based on family background in reading literacy, as evidenced by assessments like PISA and PIRLS.

In Slovenia, the system of upper secondary education is divided into three main educational programs. The academic track, represented by the program of general gymnasium, leads directly to university study and it is the most demanding four-year education course with the same curriculum for all future university students. Technical education programs offer a path to tertiary non-university study courses and provide certificates for different professions in the form of four-year programs. Vocational programs of four years or medium duration include longer and shorter programs oriented at vocations

and do not provide direct access to tertiary education. A secondary school can offer either only the program of general gymnasium or more programs and tracks of technical education or a combination of technical and vocational programs. Students choose the secondary school and program at the end of grade 9 of elementary school, according to their educational motivation and achievement at the end of compulsory schooling. A majority of students stay enrolled in the same program and school to the end of their secondary education. Following the findings in the research literature, upper secondary schools from different educational programs are expected to differ in their general school climate, formed by the motivation and characteristics of enrolled students as well as teachers and teaching.

The main aims of our study were to better understand the link between achievement and student attitudes and to discover the factors that most strongly affect achievement across secondary school student populations. Due to tracking in Slovenia, the basic hypothesis was that schools from different educational programs differ in terms of student population characteristics, attitudes, and knowledge. Consequently, specific approaches or actions would be needed to improve the situation at the school level. Given the research problem described above, the study aims to answer two fundamental research questions:

1. How do various aspects of student well-being and achievement, as well as teacher attitudes, vary across different educational programs?
2. What are the most significant predictors of student well-being in different educational programs?

Our study for the first time combines research findings for Slovenia from two different international comparative studies, focused on two different educational levels. Students' reports about their knowledge and attitudes are combined with teachers' reports on teaching the same student population. The study fills the gap in research on teaching the population of 15-year-old-students and the specific educational problem, motivation for learning, from both educational levels, students and teachers. The main research interest—not only to describe in general, but to find differences within secondary education, between different school programs in order to inform the national educational system about how to adapt policies to different programs, is very hard to find in contemporary research literature.

2 Data and methods

Ideally, the study of relations between students and teachers would use data sources from both teachers and students. Unfortunately, the PISA study does not collect teacher data. To still study the problem from both sides, we linked

the results from the two studies. To describe the problem and background of research questions, a descriptive analysis of the OECD Teaching and Learning International Survey (TALIS) data from teachers in Slovenia, collected in 2018, was used in the first step of our research. To find answers to the research questions, secondary analyses of the OECD PISA data from students in Slovenia, collected in 2022 have been applied. Both studies reported on relations between teachers and students, while TALIS also described school climate and PISA measured the achievement of students. The method (participants, instruments, and data analysis) is explained in detail in the following paragraphs.

2.1 Samples

The OECD Teaching and Learning International Survey (TALIS) in 2018 collected reports from more than 40 populations of teachers in ISCED 2 programs, and provided an optional additional collection of all data in the population of ISCED 3 program teachers. Slovenia participated in ISCED 2 and ISCED 3 modules in TALIS 2018, where the population of ISCED 3 consists exclusively of teachers in Slovene upper secondary schools. All secondary schools were included in the survey due to the fact that the total number of secondary schools in Slovenia is smaller than the internationally required school sample size. Within each school, a set of 30 teachers was sampled, or all teachers if there were fewer than 30 employed in a school. Teachers replied to the online questionnaire in the spring 2018. As stated above, the majority of schools offer more than one educational program to their students. Since teachers in a particular school can teach in different educational programs offered by the school, they are not assigned to a specific program but rather to the school itself. To compare data among different programs, for this study, all participating schools and their teachers were assigned to three general educational program groups based on which of the three programs the majority of students inside the school are enrolled in: general gymnasia, technical education programs, and vocational education programs. Statistics reveal that almost one-third of teachers come from schools providing the most advanced academic program of general gymnasia, which offers students the general Higher School Certificate examination and entry to academic university studies. A quarter of teachers teach in schools where most students are enrolled in vocational programs of shorter duration (less than 4 years). The remaining 44% of teachers work in technical secondary schools that provide different programs with vocational Higher School Certificate examinations for students, leading to non-academic university or tertiary education (Table 1).

Table 1
Distribution of teachers into three groups of schools by educational programs

Educational programs	Weighted *N*	% of teachers	% of teachers *SE*
General gymnasia	1592	31.61	0.36
Technical education	2340	44.36	0.38
Vocational education	1293	24.03	0.37

The population for the PISA study in Slovenia includes all male and female students aged between 15 years and 3 months and 16 years and 2 months, regardless of the educational program they attend. Sampling in the PISA survey is multi-level and stratified. In Slovenia, the sample includes all secondary education schools. The sample in the analysis includes a representative sample of 5,591 15-year-old male and female students, of which 2,591 (49.5%) are female and 3,000 (50.6%) are male. Of the students sampled 1,486 (34.5%) students attended a program of general gymnasia, 2,766 (48.8%) attended programs of technical education, and 1,339 (16.7%) attended vocational education programs of medium duration. In general gymnasia, there were 910 (62.0%) females and 576 (38.0%) males, in technical education programs 1,286 (47.8%) females and 1,480 (52.2%) males, and in programs of vocational education of medium duration 395 (28.2%) females and 944 (71.8%) males.

2.2 Instruments and variables

The TALIS questionnaire contained questions for measuring attitudes of teachers, asking one general question about their attitude, followed by precise questions or statements in the context of the same factor. The reported answers were internationally modelled by IRT method into scales called teacher indices. The values of indices were included in the internationally available public database of TALIS 2018. For our study, we used five teacher self-reporting scales on relations between teachers and students: *Teacher-student relations* (STUD) (e.g., agreement with teachers and students usually get on well with each other; teachers believe that the students' well-being is important; teachers are interested in what students have to say); *Social utility motivation to teach* (SOCUT) (e.g., importance of teaching to influence the development of children and young people; to benefit the socially disadvantaged; to provide a contribution to society); Index *Classroom management* (CLASM) (e.g., frequency of telling students to follow classroom rules; of telling them to listen to teachers; of calming students who are disruptive; of quietening students down at the beginning of the lessons); *Student behavior stress* (STBEH) (e.g., level of stress because of being responsible for students' achievement; maintaining classroom discipline; being intimidated or verbally abused by

students); *Teachers perceived disciplinary climate* (DISC) (e.g., to what extent do they have to wait a long time for students to quiet down; do students create a pleasant learning atmosphere; is a lot of time lost because students interrupt the lesson; is there much disruptive noise in the classroom). Detailed descriptions of scales are available in the TALIS 2018 Technical Report (OECD, 2019).

TALIS scales have a mean of 10 and a standard deviation of 2. Scores on scales greater than 10 indicate average or higher agreement with the items of the scale (such as "frequently" or "always"). Scores lower than 10 indicate average disagreement or less with the items in the scale (such as "not at all" or "to some extent"). All the scales showed good measurement characteristics for the Slovenian sample (omega statistics are 0,774 for STUD; 0.789 for SOCUT; 0.769 for STBEH and 0.878 for DISC and Cronbach's alpha for CLASM is 0.865; see Stancel-Piątak et al., 2019).

After a two-hour PISA 2022 reading, mathematics, and science literacy assessment, students completed the student questionnaire, which is used to identify the effects of different background factors on student achievement. For the analysis, we used separate student self-reported well-being scales: Quality of student-teacher relationships (QUALITY) (e.g., "The teachers at my school are respectful towards me", "The teachers at my school are interested in students' well-being"), Sense of belonging to school (BELONG) (e.g., "I feel like an outsider (or left out of things) at school", "I feel like I belong at school"), Frequency of being bullied (BULLIED) (e.g., "Other students left me out of things on purpose", "Other students took away or destroyed things that belonged to me"), and Feeling of being safe at school and outside school (FEELSAFE) (e.g., "I feel safe on my way to school", "I feel safe on my way home from school"). Students responded to 27 statements altogether[1]. All the mentioned scales and corresponding items are described in detail in the PISA 2022 Technical Report (OECD, 2023b). All the scales showed good internal consistency for the Slovenian sample, with the coefficient alpha values ranging between $\alpha = 0.74$ and $\alpha = 0.91$ (OECD, 2023b).

Mathematics, reading, and science achievement in PISA 2022 are internationally comparable average measures of mathematics, reading, and science ability (OECD, 2023a). The average mathematics achievement of Slovenian 15-year-

[1] Due to the high number of the scales and related statements, only a few examples of the statements are given for each scale in continuation of the description of used instruments. The full scales are available in the internationally published questionnaire at the following link: https://www.oecd.org/pisa/data/2022database/CY8_202111_QST_MS_STQ_CBA_NoNotes.pdf

olds was 485 points on an International Mathematics scale, and as such significantly above the OECD average (472 points). The average reading achievement of Slovenian 15-year-olds was 469 points on an International Reading scale, and as such significantly below the OECD average (476 points). The average science achievement of Slovenian 15-year-olds was 500 points on an International Science scale, and as such significantly above the OECD average (485 points) (Šterman Ivančič & Mlekuž, 2023). However, the trend from the previous PISA cycle was negative for all three measures of achievement.

2.3 Data analysis

Mean values of selected TALIS indices STUD, SOCUT, CLASM, STBEH and DISC were calculated for three groups of teachers according to the educational programs offered by their schools. Differences between the groups were observed.

For the analysis of PISA data, we used pre-existing average values of achievement in Mathematics, Reading, and Science, and standardized values of well-being indices (i.e. perceived quality of student-teacher relationships, sense of belonging to school, frequency of being bullied, feeling of being safe)[2] for Slovenia from the PISA 2022 database. First, we compared the mean values of mathematics, reading, and science achievement, and the well-being indices with the OECD average. Since we were interested in the effects of students' perceived quality of student-teacher relationships, sense of belonging to school, frequency of being bullied, and feeling of being safe at school and outside school on students' average mathematics, reading, and science achievement, we used the linear regression procedure to further analyze the size effects. Pearson correlation coefficients were used to examine associations between the independent (BELONG, BULLIED, FEELSAFE, QUALITY) and the dependent variables (MATHEMATICS, READING, SCIENCE ACHIEVEMENT) prior to undertaking regression. We also undertook preliminary screening of the data through the examination of residuals with the scatterplot of residuals against predicted values and checked the data for multivariate outliers (Tabachnick and Fidell, 2007). No obvious pattern to the errors and no multivariate outliers were identified.

---

[2] The index is a standardized value with a mean of 0 and a standard deviation of 1 and enables the comparison of Slovenian scores on a certain scale to the OECD average. Negative values of the index mean that Slovenian students on average evaluated their aspects of well-being on a certain scale lower than their peers from the OECD countries, and positive values of the index mean that their self-evaluation on a certain scale was higher.

Data from both sources underwent analysis using the statistical program IEA IDB Analyzer (Version 5.0.23). This program processes data from two-stage sampling in both studies and in the case of PISA incorporates weights for individual sample units. It also enables us to accurately assess the standard errors of estimated parameters in the population using the Bootstrap method.

3 Results

3.1 Teachers' reports about relations with students

The first part of our study focused on the background conditions in the schools. The descriptive analysis of teachers' attitudes indicates differences among these attitudes in the three groups of schools, providing more or less advanced educational programs and consequently enrolling different student populations. The means and *t*-values of comparisons between educational programs are given in the Table 2.

Table 2
Means of indices of relations between teachers and students by groups of teachers

Index	*M*	*SE*	Comparing group	*t*
STUD				
General gymnasia	13.13	0.08	Technical education	–0.88
Technical education	13.05	0.06	Vocational education	–2.36*
Vocational education	12.84	0.08	General gymnasia	2.54*
SOCUT				
General gymnasia	11.56	0.09	Technical education	3.06**
Technical education	11.94	0.07	Vocational education	–0.10
Vocational education	11.95	0.10	General gymnasia	2.93*
CLASM				
General gymnasia	10.29	0.09	Technical education	4.52***
Technical education	10.81	0.08	Vocational education	3.43***
Vocational education	11.23	0.09	General gymnasia	–7.15***
STBEH				
General gymnasia	8.65	0.08	Technical education	5.27***
Technical education	9.22	0.07	Vocational education	4.12***
Vocational education	9.71	0.09	General gymnasia	–8.85***
DISC				
General gymnasia	7.84	0.08	Technical education	3.94***
Technical education	8.24	0.07	Vocational education	4.04***
Vocational education	8.67	0.08	General gymnasia	–7.69***

Note. *$t > 1.98$, **$t > 2.68$, ***$t > 3.28$.

Results reveal expected patterns of means of indices for three groups. The teacher student relations (STUD) decrease in value from gymnasia to vocational education, while mean values of the indices social utility motivation to teach (SOCUT), classroom management (CLASM), student behavior stress (STBEH), and teachers' perceived disciplinary climate (DISC) increase from gymnasia to vocational education (Figure 1).

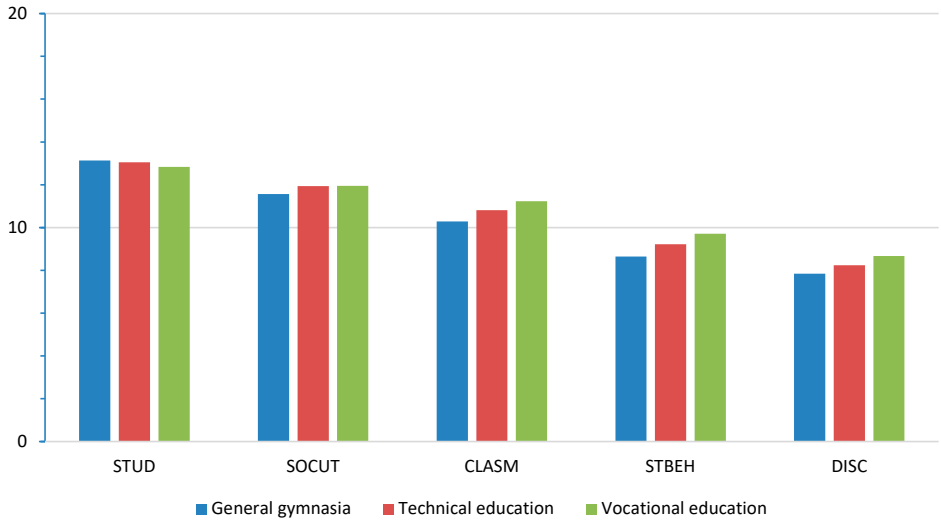


Figure 1
Means of teacher indices by educational programs

The *teacher student relations* index means are larger than the scale mean, 10 points, in all groups, showing that teachers in general agree or strongly agree with all its statements about relations. Higher index values mean better relations, so according to the results, teachers across all schools believe relations between students and teachers are good. The index differs only slightly between the groups of teachers. The *teacher student relations* are lower in vocational schools than in gymnasia and in technical schools, while they are similarly high in gymnasia and in technical schools.

Social utility motivation to teach is the lowest among teachers in gymnasia and higher in technical and vocational schools, but does not differ between them. When deciding to become teachers, teachers in gymnasia have seen the social role of teaching less as a motivating factor than their colleagues in technical and vocational schools. *The classroom management*, indicating the frequency of teachers' requests for discipline from students, shows that teachers request discipline from students in all three educational groups a

lot, frequently (or more) or in all lessons. The mean as expected increases from gymnasia through technical to vocational schools, reporting the most teacher requests for discipline in vocational schools. Teachers in these schools were also confirmed to be the most stressed because of students. The indices of *student behavior stress* and *teachers' perceived disciplinary climate* follow the same increasing pattern from gymnasia to vocational schools. Teachers in gymnasia felt the least stress and teachers in vocational schools felt the most because of being held responsible for student achievement, maintaining classroom discipline, and being intimidated or verbally abused by students. Similarly, teachers in gymnasia also reported the highest discipline in classes while in vocational schools the measured values of the index of discipline are the lowest among the three groups of schools.

In general, teachers in gymnasia reported the most positive relations with students and the fewest disciplinary issues of all three groups. Teachers in technical schools reported slightly less positive relations with students than in gymnasia, but more positive than in vocational schools. In addition, they had more disciplinary issues with students than in gymnasia but fewer than in vocational schools. The results clearly confirm differences in school climate among all three groups of schools according to their educational programs, with the most positive situation in the schools with the most advanced programs.

3.2 Student achievement

Compared to the OECD average, Slovenian 15-year-olds achieved above-average results in mathematics and science literacy in PISA 2022 (485 vs. 472 in mathematics and 500 vs. 485 in science), while in reading, the Slovenian average achievement was significantly below the OECD average (469 vs. 476).

Looking at the differences in average achievement in mathematics, reading, and science literacy by educational program (Table 3), the results show that in all three subjects, students attending gymnasia achieved the highest average scores, while students attending vocational education programs achieved the lowest.

Table 3

Average achievement in mathematics, reading, and science according to secondary educational programs

Educational programs	Mathematics		Reading		Science		
	M	*SE*	*M*	*SE*	*M*	*SE*	
General gymnasia	555.72	2.24	541.74	2.59	571.59	2.76	
Technical education	469.43	1.78	453.09	2.09	484.82	2.45	
Vocational education	407.27	2.21	383.63	3.06	419.32	3.08	

The differences in average achievement between the most and least advanced programs are quite large and significant, with 148 points in mathematical literacy, 158 points in reading literacy, and 122 points in science literacy.

3.3 Students' reports about attitudes and relations with teachers

Furthermore, compared to the OECD average, Slovenian 15-year-olds reported significantly lower-than-average quality of both student-teacher relationships and frequency of being bullied, and average levels of sense of belonging to school and feelings of being safe in school (Table 4).

Table 4
Mean standardized values of well-being indices for study variables

Student indices	*M*	*SE*
Quality of student-teacher relationships	–0.21	0.01
Sense of belonging to school	0.04	0.01
Being bullied	–0.43	0.01
Feeling safe	0.02	0.02

Note. M = Mean value of internationally comparable index – standardized value with the mean of 0 and standard deviation of 1.

Again, there are significant differences in well-being indices between educational programs (Table 5). Students attending vocational education programs of medium duration report significantly lower quality of student-teacher relationships, a lower sense of belonging to school and feelings of being safe, and slightly higher exposure to peer violence compared to students attending general gymnasium.

Table 5
Mean values of well-being indices for study variables by educational programs

Educational programs	Quality of relationships		Sense of belonging		Being bullied		Feeling safe	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
General gymnasia	–0.08	0.03	0.08	0.03	–0.58	0.02	0.17	0.03
Technical education	–0.30	0.02	0.04	0.02	–0.43	0.02	–0.06	0.02
Vocational education	–0.30	0.03	–0.06	0.03	–0.33	0.04	–0.12	0.04

Note. M = Mean value of internationally comparable index – standardized value with the mean of 0 and standard deviation of 1 (except for average mathematics, reading, and science achievement).

All variables are significantly correlated with each other at the $\alpha < 0.001$ level (Table 6), with the quality of student-teacher relationships being most strongly associated with mathematics, reading, and science achievement.

Table 6
Correlation coefficients for study variables

Study variables	1	2	3	4	5	6	*M*	*SE*
Quality of student-teacher relationships	–						–0.21	0.01
Sense of belonging to school	0.26**	–					0.04	0.01
Being bullied	–0.22**	–0.27**	–				–0.43	0.01
Feeling safe	0.20**	0.36**	–0.15**	–			0.02	0.02
Mathematics achievement	0.14**	0.06**	–0.10**	0.14**	–		485	1.20
Reading achievement	0.18**	0.07**	–0.15**	0.13**	0.77**	–	469	1.60
Science achievement	0.16**	0.06**	–0.11**	0.14**	0.89**	0.76**	500	1.40

Note. The Pearson correlation coefficient was calculated as a measure of correlation; $^{**}p < 0.001$.

In explaining mathematics achievement (Table 7), the main significant predictors were the feeling of safety at school and the quality of student-teacher relationships, both in the general gymnasium and the technical education program. In the latter, the effect size of the quality of student-teacher relationships in explaining mathematics achievement is the largest. In the vocational education program of medium duration, none of the predictors was found to be significant in predicting mathematics achievement. The explanatory power of the model is low within all three educational programs ($R^2 = 0.03$ for general gymnasium, 0.02 for technical education programs, and 0.00 for vocational education programs of medium duration).

Table 7
Regression of association between different aspects of student well-being and mathematics achievement in three types of educational programs

Mathematics achievement					
Model	*B*	SE_B	β	t_B	*p*
General gymnasia					
Constant	554.50	2.88		192.72	0.00
Belong	–6.30	3.11	–0.08	–2.03	0.05
Bullied	–2.81	2.89	–0.03	–0.97	>0.05
Feel safe	9.79	2.61	0.13	3.75	0.00
Quality	9.14	2.33	0.12	3.92	0.00
Technical education					
Constant	471.32	2.11		223.40	0.00
Belong	–0.07	2.46	0.00	–0.03	>0.05
Bullied	3.05	1.76	–0.04	–1.74	>0.05
Feel safe	5.08	2.05	0.07	2.48	0.05
Quality	5.39	2.04	0.07	2.64	0.01
Vocational education					
Constant	410.98	2.58		159.23	0.00
Belong	0.72	2.39	0.00	0.30	>0.05
Bullied	–1.04	2.13	–0.04	–0.49	>0.05
Feel safe	–0.49	2.40	0.07	–0.20	>0.05
Quality	3.35	2.29	0.07	1.46	>0.05

Note. R^2 adjusted = 0.03 for general gymnasium; 0.02 for technical education programs; 0.00 for vocational education programs of medium duration.

In predicting achievement in reading (Table 8), the main predictors confirmed as significant across the three educational programs were the frequency of experiencing peer violence and the quality of student-teacher relationships, with the latter predictor having the largest effect in explaining achievement in reading amongst all three literacies. Again, the explanatory power of the model is low, but the highest of all three literacies (R^2 = 0.04 for general gymnasium, 0.03 for technical education programs, and 0.05 for vocational education programs of medium duration).

Table 8
Regression of association between different aspects of student well-being and reading achievement in three types of educational programs

Reading achievement					
Model	*B*	SE_B	β	t_B	*p*
General gymnasia					
Constant	538.24	3.10		173.59	0.00
Belong	–5.38	3.39	–0.07	–1.72	>0.05
Bullied	–8.24	2.87	–0.10	–2.87	0.01
Feel safe	8.19	2.74	0.10	2.99	0.01
Quality	10.26	3.32	0.13	3.08	0.01
Technical education					
Constant	456.23	2.44		187.30	0.00
Belong	–0.73	2.49	–0.01	–0.29	>0.05
Bullied	–6.64	1.96	–0.09	–3.38	0.00
Feel safe	2.29	2.31	0.03	0.99	>0.05
Quality	10.96	2.43	0.13	4.50	0.00
Vocational education					
Constant	389.83	3.33		117.11	0.00
Belong	3.99	3.01	0.05	1.33	>0.05
Bullied	–4.85	2.42	–0.08	–2.00	0.05
Feel safe	–1.76	3.12	–0.02	–0.56	>0.05
Quality	12.50	3.14	0.17	3.98	0.00

Note. R^2adjusted = 0.04 for general gymnasium; 0.03 for technical education programs; 0.05 for vocational education programs of medium duration.

In explaining achievement in science (Table 9), the quality of student-teacher relationships was also confirmed as a significant predictor of science achievement across all three educational programs, with a sense of safety at school also confirmed as a significant predictor in the general gymnasia and vocational education programs. The highest portion of the variability in science achievement is explained within the general gymnasium program, but the explanatory power of the model is low (R^2 = 0.04 for general gymnasia, 0.02 for technical education programs, and 0.02 for vocational education).

Table 9
Regression of association between different aspects of student well-being and science achievement in three types of educational programs

Science achievement					
Model	*B*	SE_B	β	t_B	*p*
General gymnasia					
Constant	570.83	3.07		185.64	0.00
Belong	–4.95	3.16	–0.06	–1.57	>0.05
Bullied	–2.79	2.76	–0.03	–1.01	>0.05
Feel safe	9.66	3.00	0.12	3.21	0.00
Quality	10.56	3.02	0.13	3.50	0.00
Technical education					
Constant	487.86	2.57		189.64	0.00
Belong	–0.88	3.17	–0.01	–0.28	>0.05
Bullied	–3.07	2.01	–0.04	–1.52	>0.05
Feel safe	5.14	2.60	0.06	1.97	0.05
Quality	8.08	2.28	0.10	3.54	0.00
Vocational education					
Constant	423.63	3.21		131.84	0.00
Belong	–0.01	2.91	0.00	0.00	>0.05
Bullied	–5.30	2.41	–0.09	–2.20	0.05
Feel safe	0.73	2.89	0.01	0.25	>0.05
Quality	6.59	2.52	0.09	2.62	0.01

Note. R^2adjusted = 0.04 for general gymnasium; 0.02 for technical education programs; 0.02 for vocational education programs of medium duration.

4 Discussion

The teacher-student relations as measured through teachers' responses appear high (above international average) with differences across educational programs. At the same time students report somewhat more critically, their mean values in all programs are below international averages. This discrepancy allows no absolute conclusion and might even be an inherent result of using questionnaire reports as the only data. The quality of teacher-student relationships therefore remains open to different interpretations, which is a general limitation of our data. Answers to the first research question, how do various aspects of student well-being and achievement, as well as teacher attitudes, vary across different educational programs vary across educational tracks. Lower results in technical and vocational educational programs can be attributed to differences in student composition. Previous research in the

Slovenian educational system (Cankar & Zupanc, 2020) demonstrated that when students progress to upper secondary education, the most demanding educational track (general gymnasium) tends to attract a more homogenous population with more students with higher social-economic status and fewer students with disabilities, learning difficulties or disadvantages. Consequently, other educational programs receive a more diverse student population leading to more demanding teacher student relations and lower scores on the questionnaire. This is supported by other indices with teachers reporting more classroom management, more stress from student behavior and worse perceived disciplinary climate in vocational and technical educational programs. This is consistent with students' self-reported data on well-being – all indices show the same pattern of differences between educational programs.

The first results of the analysis of PISA data on students' self-reported well-being in Slovenia show that, compared to their peers in OECD countries, Slovenian students reported somewhat lower levels of different aspects of their well-being. This is especially evident in their perceived quality of student-teacher relationships, feelings of belonging to the school, and feelings of safety at school. Notably, this trend is particularly pronounced in technical education and vocational education programs. Moreover, the results from TALIS also confirm that the relationships between students and teachers, as reported by teachers, are at least slightly better in schools offering more advanced programs. These findings align with the fact that these schools enroll higher-achieving students who are better motivated to learn and wish to continue further education at universities. Additionally, the indices measuring the amount of teacher effort to maintain discipline and engage students to manage a positive class climate increase from the most academically demanding program to the least demanding one. Again, one of the reasons could be that students in gymnasia are more motivated for school work than in vocational programs. Additional reasons for increased reports of teachers' efforts to maintain an orderly class climate in vocational schools can be also a consequence of the nature of vocational programs, which include many practical subjects where students are asked to actively work with materials and devices or in teams during lessons, challenges for a quiet and orderly climate. In general, according to teachers' reports, the relations and climate are better in gymnasia than in technical and vocational education. In the view of students, the relations are also better in gymnasia than in vocational educational programs. We may therefore conclude that there are differences in indices of different aspects of well-being between teachers and students in different educational programs. The finding is supported by results of other studies of student well-being in different tracks, academic and vocational, showing the differences in a variety of factors, such as self-concepts and sense

of purpose (Jónsdóttir & Blöndal, 2022), sense of belonging, teachers' trust in students (Van Houtte & Van Maele, 2012), self-esteem and school belonging (Wu & Becker, 2023), mostly in favor of students in academic tracks, but also the crucial role of teacher support to strengthen students' attitudes of well-being.

The study results also show the answer to the second research question, what are the most significant predictors of student well-being in different educational programs. Although the variance of student achievement in mathematics, reading, and science that can be explained by different aspects of students' well-being is relatively low, there is one predictor that proved to be significant across all three domains and in all three educational programs, namely the quality of student-teacher relationships, where the most significant effects were observed when explaining student achievement in reading.

The results of the study therefore suggest that teachers play a crucial role in shaping students' academic performance and psychological well-being through their behavior and interactions within the classroom, and are in line with the studies (e.g., Prewett et al., 2019) that highlight that teachers' prosocial classroom behavior and social-emotional support are the strongest predictors of students' perceptions of high-quality relationships with their teachers. This finding is also supported by results from the Hand in Hand interventions program (Kozina et al., 2020) where the results showed that strengthening the social-emotional skills of both teachers and students has positive effects on various aspects of the learning process, and that a learning environment in which the student perceives that the teacher is working with them to achieve better results has a positive effect, especially on low-achieving students. This finding is particularly important in view of the differences in student well-being within different educational programs in Slovenia. Although strengthening social-emotional skills would empower teachers in any educational program, this research clearly showed that teacher-student relations in technical and vocational educational programs are most strenuous and teachers in these programs would benefit most from such support.

We can conclude that a better understanding of the interplay between factors of school climate and academic outcomes is essential for informing evidence-based interventions aimed at promoting student success and well-being in educational settings. By empowering teachers and influencing classroom practices, educators can create supportive learning environments that empower students to overcome challenges, strive for continuous improvement, and achieve academic success, ultimately promoting both their performance and well-being.

Our study has limitations. The samples from both studies were not designed in advance to be directly linked. The sample of students for PISA include only 15-year-old students, most of them in the first grade of their

secondary schools. The sample of teacher for TALIS includes all teachers of all grades in secondary schools and questionnaires collected their reports on teaching all grades, not only the student population represented by the PISA sample. The interpretations of results are therefore based on more general teacher opinions than on teaching only PISA students. In addition, we are aware of specific differences in student characteristics across school programs that could be related with differences in student opinions, such as gender, SES or educational expectations. Although student populations in general secondary schools are mostly not gender biased, programs in some vocational schools do enroll mostly boys or girls. Unfortunately, available samples for our study statistically did not allow us the extensions to search for differences in vocational programs by student gender or detailed study of impact of student SES on student opinions or teachers' reports. With results showing the need for specific attention to the vocational programs, further studies and more detailed data are needed to address these issues.

References

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review*, *4*(3), 165–176. https://doi.org/10.1016/j.edurev.2009.04.002

Bosworth, K., Ford, L., & Hernandaz, D. (2011). School climate factors contributing to student and faculty perceptions of safety in select Arizona schools. *Journal of School Health*, *81*(4), 194–201. https://doi.org/10.1111/j.1746-1561.2010.00579.x

Braun, H. (2008). Review of McKinsey report: How the world's best performing school systems come out on top. *Journal of Educational Change*, *9*(3), 317–320. https://doi.org/10.1007/s10833-008-9075-9

Burris, J. E. (2012). It's the teachers. *Science*, *335*(6065), 146–146. https://doi.org/10.1126/science.1218159

Cankar, G., & Zupanc, D. (2020). *Pravične možnosti izobraževanja*. Državni izpitni center.

Carnoy, M., Khavenson, T., Loyalka, P., Schmidt, W. H., & Zakharov, A. (2016). Revisiting the relationship between international assessment outcomes and educational production: Evidence from a longitudinal PISA-TIMSS sample. *American Educational Research Journal*, *53*(4), 1054–1085. https://doi.org/10.3102/0002831216653180

Contini, D., & Cugnata, F. (2020). Does early tracking affect learning inequalities? Revisiting difference-in-differences modeling strategies with international assessments. *Large-Scale Assessments in Education*, *8*(1), 14. https://doi.org/10.1186/s40536-020-00094-x

Eklöf, H., Pavešič, B. J., & Grønmo, L. S. (2014). A cross-national comparison of reported effort and mathematics performance in TIMSS advanced. *Applied Measurement in Education*, *27*(1), 31–45.
https://doi.org/10.1080/08957347.2013.853070

Gläser-Zikuda, M., & Fuß, S. (2008). Impact of teacher competencies on student emotions: A multi-method approach. *International Journal of Educational Research*, *47*(2), 136–147.
https://doi.org/10.1016/j.ijer.2007.11.013

Goodenow, C., & Grady, K. E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *The Journal of Experimental Education*, *62*(1), 60–71.
https://doi.org/10.1080/00220973.1993.9943831

Hagenauer, G., & Volet, S. E. (2014). Teacher–student relationship at university: An important yet under-researched field. *Oxford Review of Education*, *40*(3), 370–388.
https://doi.org/10.1080/03054985.2014.921613

Hughes, J. N. (2011). Longitudinal effects of teacher and student perceptions of teacher-student relationship qualities on academic adjustment. *The Elementary School Journal, 112*(1), 38–60.
https://doi.org/10.1086/660686

Jónsdóttir, H. H., & Blöndal, K. S. (2022). The choice of track matters: Academic self-concept and sense of purpose in vocational and academic tracks. *Scandinavian Journal of Educational Research, 67*(4), 621–636.
https://doi.org/10.1080/00313831.2022.2042843

Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., & De Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education*, *35*(6), 641–680.
https://doi.org/10.1080/02671522.2019.1615116

Kozina, A. (Ed). (2020). *Social, emotional and intercultural competencies for inclusive school environments across Europe: Relationships matter.* Studien zur Schulpädagogik, Band 89. Dr. Kovač.

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, *65*, 50–64.
https://doi.org/10.1016/j.lindif.2018.05.009

McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student–teacher relationships and their outcomes. *Educational Research Review*, *14*, 1–17.
https://doi.org/10.1016/j.edurev.2014.12.001

Mikk, J., Krips, H., Säälik, Ü., & Kalk, K. (2016). Relationships between student perception of teacher-student relations and PISA results in mathematics and science. *International Journal of Science and Mathematics Education*, *14*(8), 1437–1454.
https://doi.org/10.1007/s10763-015-9669-7

OECD. (2023a). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. OECD.
https://doi.org/10.1787/53f23881-en

OECD (2023b). *PISA 2022 Technical report*.
https://www.oecd.org/pisa/data/pisa2022technicalreport/.

OECD (2019). *TALIS 2018 Technical Report*.
https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/talis/migrate/TALIS_2018_Technical_Report.pdf

Papanastasiou, E. C., Zembylas, M., & Vrasidas, C. (2003). Can computer use hurt science achievement? The USA results from PISA. *Journal of Science Education and Technology*, *12*(3), 325–332.
https://doi.org/10.1023/A:1025093225753

Petko, D., Cantieni, A., & Prasse, D. (2017). Perceived quality of educational technology matters: a secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *Journal of Educational Computing Research*, *54*(8), 1070–1091.
https://doi.org/10.1177/0735633116649373

Prewett, S. L., Bergin, D. A., & Huang, F. L. (2019). Student and teacher perceptions on student-teacher relationship quality: A middle school perspective. *School Psychology International*, *40*(1), 66–87.
https://doi.org/10.1177/0143034318807743

Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, *88*(3), 408–422.
https://doi.org/10.1037/0022-0663.88.3.408

Sari, M. (2012). Sense of school belonging among elementary school students. *C.U. Faculty of Education Journal, 41*(1), 1–11.
https://dergipark.org.tr/en/pub/cuefd/issue/4133/54251

Siann, G., Callaghan, M., Glissov, P., Lockhart, R., & Rawson, L. (1994). Who gets bullied? The effect of school, gender and ethnic group. *Educational Research*, *36*(2), 123–134.
https://doi.org/10.1080/0013188940360202

Stancel-Piątak, A., Wild, J., Chen, M., Rozman, M., Mirazchiyski, P. & Cigler H. (2019). Validation of scales and construction of scale scores. In OECD, *TALIS 2018 Technical Report* (pp. 191–433). OECD.
https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/talis/migrate/TALIS_2018_Technical_Report.pdf

Strello, A., Strietholt, R., Steinmann, I., & Siepmann, C. (2021). Early tracking and different types of inequalities in achievement: Difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability*, *33*(1), 139–167.
https://doi.org/10.1007/s11092-020-09346-4

Šterman Ivančič, K. & Mlekuž, A. (2023). *PISA2022: program mednarodne primerjave dosežkov učencev in učenk: nacionalno poročilo s primeri nalog iz matematike*. Educational Research Institute.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Pearson Education Inc. and Allyn & Bacon.
https://doi.org/10.1007/s11205-007-9093-7

Van Petegem, K., Aelterman, A., Van Keer, H., & Rosseel, Y. (2007). The influence of student characteristics and interpersonal teacher behaviour in the classroom on student's wellbeing. *Social Indicators Research*, *85*(2), 279–291.

Van Houtte, M., & Van Maele, D. (2012). Students' sense of belonging in technical/vocational schools versus academic schools: The mediating role of faculty trust in students. *Teachers College Record*, *114*(7), 1–36. https://doi.org/10.1177/016146811211400706

Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying? Recommendations to enhance the validity, transparency, and comparability of bullying research. *Aggression and Violent Behavior*, *36,* 34–43. https://doi.org/10.1016/j.avb.2017.07.003

Williams, S., Schneider, M., Wornell, C., & Langhinrichsen-Rohling, J. (2018). Student's Perceptions of School Safety: It Is Not Just About Being Bullied. *The Journal of School Nursing*, *34*(4), 319–330. https://doi.org/10.1177/1059840518761792

Wu, Y.-J., & Becker, M. (2023). Association between school contexts and the development of subjective well-being during adolescence: A context-sensitive longitudinal study of life satisfaction and school satisfaction. *Journal of Youth and Adolescence, 52*(5), 1039–1057. https://doi.org/10.1007/s10964-022-01727-w

STUDIA PAEDAGOGICA

vol. 29, n. 3, 2024
https://doi.org/10.5817/SP2024-3-2

STUDY

IF THEY TALK MORE DURING LESSONS, WILL THEY ACHIEVE BETTER? UNLOCKING THE RECIPROCAL RELATIONSHIP BETWEEN STUDENT VERBAL PARTICIPATION AND ACHIEVEMENT

Martin Sedláček[a], Klára Šeďová[a], Roman Švaříček[a], Zuzana Šalamounová[a]

[a]Faculty of Arts, Masaryk University

ABSTRACT

This study investigates the relationship between student verbal participation and achievement in sixth-grade language arts. We conducted an intervention in six classes to enhance and equalize student talk while reducing individual disparities in participation. The design of the study involved measuring talk time and administering reading literacy tests before and after the intervention, with similar measurements in six control classrooms. The results indicated increased and more evenly distributed verbal participation in the intervention classrooms than in the control classrooms. However, no significant differences in student achievement were observed between the two groups. A path analysis examined the link between participation and achievement, confirming that verbal participation is a predictor of student success. The study suggests that the impact of increased verbal participation on achievement might be more pronounced over the long term, necessitating further research with delayed post-measurements to fully understand this relationship.

KEY WORDS

student talk; student achievement; reading literacy test; intervention study; path analysis; post covid study

CORRESPONDING AUTHOR
Martin Sedláček, Faculty of Arts, Masaryk University, Arna Nováka 1/1, 602 00 Brno, Czech Republic
e-mail: msedlace@phil.muni.cz

Introduction

The question of how to provide high-quality instruction for all students is central in the field of educational sciences. It is believed that actions taken within the school environment exert a considerable influence on the disparities in student academic performances within various social and economic settings (Morlà-Folch et al., 2022). Research into how to help all students improve their educational outcomes regardless of their socioeconomic background is very much alive, with mounting evidence indicating that factors at the classroom level have a greater capacity to account for variations in student achievement than factors at other levels (Panayiotou et al., 2021; Kyriakides et al., 2020).

Research at the classroom level often emphasizes the effective instructional practices of teachers; the understanding of student roles in instructional processes is less emphasized (Schenke, 2018). However, it has been credibly confirmed that student engagement and participation matter and are decisive for student outcomes (Bae & DeBusk-Lane, 2019; Chang et al., 2016; Decristan et al., 2023; Schnitzler et al., 2020). In this study, we view teacher and student practices as complementing each other in co-constructing the quality of instruction in the classroom.

Classroom talk during whole class teaching and the role of students within it is a primary focus of this study. It has been repeatedly found that classroom dialogue matters for students and that optimal patterns of classroom talk can enhance student achievement (Alexander, 2018; Hardman, 2016; Howe et al., 2019). Additionally, studies have shown that individual students in the class benefit differently from classroom dialogue depending on how intensively they participate in classroom discourse (Šeďová et al., 2019; Decristan et al., 2023; Ing et al., 2015; Neuman et al., 2021; Rüede et al., 2023; Schnitzler et al., 2020; Webb et al., 2014). Based on the findings from these studies, we closely examine the link between student verbal participation and student achievement. We want to confirm the existence of the link and to explore the potential for utilizing the link to enhance student learning and performance. We conducted an intervention project aimed at equalizing verbal participation among sixth-grade students. By analyzing data on student verbal participation in language arts lessons and their performance in literacy tests, we aim to determine whether it is possible to influence student outcomes through an increase in their verbal participation.

1 Individual differences in student participation in classroom discourse

Student participation in classroom discourse has been extensively studied in the past decade. It has been operationalized through students' verbal contributions, including the frequency and length of those contributions (Clarke, 2015; Decristan et al., 2023; Helgevold, 2016; Jurik et al., 2013; O'Connor et al., 2017; Šeďová et al., 2019), as well as various forms of signaling student intent to participate, such as hand-raising and calling out (Böheim et al., 2020; Decristan et al., 2023; Mundelsee & Jurkowski, 2021; Orner & Netz, 2023; Schnitzler et al., 2020). A unanimous finding across these studies is the wide variation in participation among individual students in the classroom. Some students are vocal and keen to take the floor; others are less engaged. A significant group of students remains completely non-participatory. A recent large study from Germany (Decristan et al., 2023) found that approximately 30% of students neither spoke nor raised their hands during observed lessons.

Student participation in classroom talk is associated with several factors. Those who participate more in class have been found to have higher socioeconomic status (Kelly, 2008; Orner & Netz, 2023; Šeďová & Sedláček, 2023), to be extroverted (Caspi et al., 2006; Young, 2014), to be motivated to learn (Böheim et al., 2020; Jurik et al., 2013), and to have a high academic self-concept (Kawabe et al., 2014; Schnitzler et al., 2020). The most recurrent finding is that high-participating students are those who achieve higher grades and have more prior knowledge (Clarke, 2015; Decristan et al., 2023; Myhill, 2002; Kelly, 2008; Jurik et al., 2013; Šeďová et al., 2019). Altogether, this stream of research portrays the image of a "good student" who possesses sufficient cultural and social capital and therefore seizes the opportunity to participate in class discussions.

2 Effects of participation for student learning

Classroom discourse has been demonstrated to have a positive impact on student outcomes. Research studies have shown that students who actively participate in classroom discourse tend to achieve greater learning gains. Webb et al. (2014) and Ing et al. (2015) found that students who frequently developed their own ideas in conversations and explained them to others during math lessons achieved better results in math tests. Šeďová et al. (2019) discovered that students with higher talk time and more utterances with argument time during language arts lessons performed better in reading literacy tests. Neuman et al. (2021) conducted a study in which they found

that the number of conversational turns during a lesson predicted statistically significant improvements in young children's vocabulary scores. Schnitzler et al. (2020) discovered that students who raised their hands more frequently achieved better results in end-of-year assessments. Decristan et al. (2023) found that students who raised their hands and actively participated in discussions during math and science lessons achieved better scores in math and science tests. According to a study by Rüede et al. (2023), the number of productive discourse moves initiated by students during math lessons showed a positive correlation with their performance on math tests.

All these results are impressive, but it cannot be overlooked that the operationalization of participation differs heavily across these studies. Further, there have been studies that did not confirm this link between participation and achievement: Inagaki et al. (1998) and O'Connor et al. (2017). Therefore, exploring the link between participation and achievement still deserves systematic focus.

Another underexplored aspect of the cited studies is the problem of causality. The studies confirming that those who participate more learn more did not control for the possibility that the link has the reverse direction. Some researchers have provided evidence that high achievers and students with prior knowledge participate more (Clarke, 2015; Decristan et al., 2023; Myhill, 2002; Kelly, 2008; Jurik et al., 2013; Šeďová et al., 2019). Therefore, the link between participation and achievement must be carefully examined with both causal possibilities considered.

3 Equitable participation as a tool for enhancing achievement of all students

Given what is known about the positive link between individual participation in classroom discourse and student achievement (Decristan et al., 2023; Ing et al., 2015; Neuman et al., 2021; Rüede et al., 2023; Schnitzler et al., 2020; Šeďová et al., 2019; Webb et al., 2014), it is necessary to address who is given the opportunity to talk and be heard in class. In this sense, Vrikki et al. (2019) called for equitable participation, which they understood as ensuring that all students have equal opportunities to engage in and contribute to classroom discourse. Similarly, we advocate collective classroom dialogue as a tool for equalizing learning opportunities for students and mitigating educational inequalities resulting from student socioeconomic backgrounds (see Šeďová et al., 2019; Šeďová & Sedláček, 2023).

In this light, it is surprising intervention studies striving to enhance the participation of all students are scarce. To our knowledge, there have been only two: Sedláček & Šeďová (2020) and Moser et al. (2022). In both studies,

a professional development program was conducted that focused on improving classroom dialogue. Researchers, apart from other measurements, observed the number of students verbally participating in post-intervention lessons. Both studies were only partially successful in increasing the number of participating students, and neither study controlled for changes in student achievement.

This creates a significant research gap, which we aim to address in this study. We want to determine whether the increase in student verbal participation after the intervention will be followed by an improvement in student achievement. Additionally, we want to examine the link between participation and achievement by investigating which of these variables serves as the predictor for the other.

4 Methodology

4.1 Research design

The research design was a quasi-experiment, as a random selection of intervention and control classes was unfeasible. Our study aims to assess the impact of intervention to increase student verbal participation in language arts lessons on reading literacy achievement. We designated specific classes as "intervention" classes; their selection was contingent upon the willingness of the respective schools and teachers to participate in the intervention program. In total, six intervention classes from four mainstream schools in the South Moravian region were recruited for this study. All schools were characterized by mainstream curricula and the absence of tracking practices. It is noteworthy that all schools in our study received a "good" rating from the Czech School Inspectorate, with none falling into the categories of "excellent" or "below average." The schools with "control" classes were randomly selected from a predefined list of schools within the same region, sharing similar ratings and student enrolment figures. Data collection in both the intervention and control groups of classes was based on observations and video recordings in Czech language lessons and language arts lessons.

4.2 Intervention program and sample

The intervention was designed for Czech language teachers, with all activities subsequently implemented in their classrooms. The program included five group workshops for teachers, collaborative lesson planning conducted within teacher-researcher pairs (n = 5), video recordings of lessons during the intervention (n = 5), and video-stimulated reflections on these lessons within the teacher-researcher pairs (n = 5). Their training focused on the implementation of specific teacher talk moves, including encouraging student

ideas, facilitating students building upon these ideas, promoting reasoning, extending ideas, and posing challenging questions. These practices were inspired by the T-SEDA framework (https://www.educ.cam.ac.uk/research/programmes/tseda). Teachers were trained to focus talk moves on individual students and to be sensitive to giving support to students who stay silent or hesitate to participate. Throughout the workshop sessions, a central emphasis was placed on creating a safe and supportive classroom environment, as well as acquainting students with the importance of active listening and respecting their peers' contributions. In addition to their workshop attendance, from November 2021 to May 2022, all participating teachers collaborated in pairs with researchers. Their task involved preparing five consecutive lessons, each incorporating support of equitable participation. Before each lesson, a planning session was conducted by the teacher-researcher pairs to outline the instructional approach. Subsequently, the lessons were video recorded by the researcher. Following each lesson, a reflective session was held within the teacher-researcher pairs. During these sessions, both teacher and researcher jointly reviewed video clips from the lesson, focusing on how equitable student participation was and how the teacher supported it. Notably, these sessions provided talk time measurements for all students present during the video-recorded lesson. This allowed teachers to gain a comprehensive overview of their success in involving all students in the classroom discourse.

As mentioned above, the sample consisted of six intervention classes and six control classes. All classrooms were sixth grade, with students aged eleven to twelve years. A total of 276 students participated in the study: 145 in intervention and 131 in control classes. There were no significant differences in composition (gender, native language, socioeconomic status) between the class groups (Table 1).

Table 1
Characteristics of the sample

Intervention classes	*N* (%)
Girls	77 (53.0%)
Boys	68 (47.0%)
Other native language	7
ESeCa (working class)	20
Control classes	
Girls	57 (44.0%)
Boys	74 (56.0%)
Other native language	5
ESeCa (working class)	26

4.3 Measures

4.3.1 Talk time

We measured the quantity of talk time in the classroom, i.e., the aggregate amount of time for which a particular student spoke during the lesson in the whole class conversation related to the curriculum. The individual student talk time was calculated as the average in seconds from the two lessons taught before the start of the program ("talk time 1") and the two lessons taught after the end of the program ("talk time 2"). Teachers were instructed to teach the lessons in their usual way. All lessons lasted 45 min (2,700 s). The student talk time was not calculated from the total time of the lesson. We excluded all situations unrelated to the curriculum of the lesson, such as organizational issues and classroom management. We also excluded the parts of the lessons in which whole-class conversation did not take place, such as during individual or group student work. We then excluded the times when the whole class was reading aloud. This left us with the time dedicated to the curriculum that was relevant for whole-class conversation – the average time was 21 min (1,260 s) per lesson. Of this amount, teachers averaged about 16 min (967 s) and the students a total of 5 min (300 s) per lesson.

4.3.2 Achievement in reading literacy

The literacy tests used in this research were developed by Scio, a company that provides a system of national comparative exams for schools in the Czech Republic. We employed two versions of standardized tests, hereafter referred to as "achievement 1" and "achievement 2". These tests contained identical types and numbers of tasks, including distinguishing fact from opinion in a reading text, recognizing manipulative communication, formulating the main idea of a text, and organizing information in a text with respect to its purpose. Specific task examples are provided in Annex 1. The tests did not feature the same anchor tasks, and students completed the tests with a seven-month gap between them. The test comparability was ensured by maintaining consistent difficulty and sensitivity levels for the tasks across both test versions. Task selection was based on Scio's task database for national comparison tests, allowing for the utilization of psychometric properties from previous waves of national comparison tests for the same age group.

4.4 Data analyses

In this paper, we examine two main questions: 1. Has the intervention program been successful in influencing achievement in reading literacy? 2. What is the relationship between talk time and achievement in reading literacy, i.e., is talk time a predictor of achievement, or is it the other way around?

Statistical analyses consisted of independent samples t-test to verify differences in the scores of the reading literacy test between intervention and

control classes. A paired samples test was used to compare the differences in test results in the first and second wave of testing ("achievement 1" and "achievement 2"). We applied analysis of covariance (ANCOVA) to assess the impact of an intervention while accounting for pre-existing differences among students. The covariates were the results before the intervention ("achievement 1") and the change in talk time. We tested two models assuming different relationships between talk time and achievement with path analysis using IBM SPSS AMOS 29. We posited that "talk time 1" positively influences performance in reading literacy assessments ("achievement 1"). Subsequently, "achievement 2" is explained by "talk time 2" and "achievement 1". In Model 2, we posited the opposite direction: students with better achievement talk more frequently in instructional communication. The model fit was examined using the chi-squared test ($\chi 2$) and its degree of freedom, the Root Mean Square Error of Approximation (RMSEA; values of 0.08 or less), the Comparative Fit Index (CFI; values greater than or equal to 0.95), and the Standardized Root Mean Square Residuals (SRMR; values of 0.08 or less) (Hu & Bentler, 1999).

4.5 Research ethics

We first sought oral consent from the school principals and all the teachers to allow us to conduct the research in their schools and classrooms. In the next step, we sought written consent from the teachers and afterward we asked for the written consent of all the parents of the students participating in the observed classes. Participants were assured of confidentiality and of the ability to withdraw at any time. Five parents in the intervention classes and 13 parents in the control classes decided not to agree with their children's involvement in the research. These students were present in the class during the recorded lessons but they sat outside of the camera's frame of vision and their talk time was not measured.

All participants were assigned numbers, and any personally identifying information was removed from the data prior to analysis.

5 Results

5.1 Impact of talk time on achievement in reading literacy

The primary goal of the intervention program was to enhance talk time within the class and at the same time to equalize talk time among individual students. The intervention program was successful in this regard. The analysis was based on comparisons of the individual student talk time before the program ("talk time 1") and after the end of the program ("talk time 2"). In the intervention classes, "talk time 1" and "talk time 2" exhibited a significant

difference, with an increase in "talk time 2" (Wilcoxon signed ranks T, $p < 0.001$). Conversely, no significant increase was observed in the control classes (p=0.12). A notable effect of the intervention was the reduction in individual differences in "talk time 1" compared to "talk time 2". This reduction is evidenced by a decrease in the coefficient of variation of approximately 30%. In this paper, we investigate whether changes in individual students' talk time correlate with their improvements in reading literacy, as detailed in Table 2.

Table 2
Descriptive statistics for talk time and reading literacy test

Measure	*N* (valid)	Min	Max	Mean		*SD*
				Statistic	*SE*	
Intervention classes						
talk time 1	124	0.00	85.02	11.87	1.29	14.55
talk time 2	138	0.17	98.48	17.45	1.43	16.82
achievement 1	130	3.85	96.15	52.11	1.73	19.82
achievement 2	125	19.23	100	59.55	1.64	18.33
Control classes						
talk time 1	115	0.00	103.18	12.95	1.57	16.89
talk time 2	126	0.00	111.85	14.38	1.54	17.23
achievement 1	100	3.85	92.31	47.75	1.91	19.11
achievement 2	112	11.54	96.15	52.82	1.81	19.14

Note: Talk time – individual student talk time (average per one lesson in seconds); achievement – reading literacy (success in %).

Comparing "achievement 1" and "achievement 2" in the intervention and control classes is essential. It is important to note there were no significant differences in the composition of the intervention and control classes in terms of gender, socioeconomic status, or ethnicity (as indicated by the number of students with a different native language). Before the intervention (i.e., "achievement 1"), the average success rate was higher in the intervention classes, with a mean percentage success rate approximately 5% higher than in the control classes (approximately 48% and 52%). However, with equal variance in both class groups, this difference was not statistically significant (ANOVA, $F = 3.707$, $p > 0.05$). A substantial shift in "achievement 2" was observed in both class groups, with significance (Paired Samples Test, $t = 5.05$, $p < 0.01$; $t = 2.81$, $p < 0.05$). This growth was expected, due to student maturation and the effects of schooling. Comparing the means in the

intervention and control classes, as in the previous wave, pupils in the intervention classes achieved better results. The difference was approximately 7% (approximately 53% and 60%), and it was statistically significant (ANOVA, $F = 7.465$, $p < 0.05$).

Comparing the mean success rate in the control and intervention classes indicates that the intervention indeed affected reading literacy achievement. This analysis has a limitation: it only compares achievement in these two groups without considering changes in talk time. To address this, differences in "achievement 2" literacy between the control and intervention classes are further confirmed through a two-way ANCOVA. The independent factor is the intervention, and the model includes a covariate represented by "talk time 2" along with "achievement 1" as a control covariate. ANCOVA offers a more refined estimate of the intervention's impact on "achievement 2" by eliminating the influence of other variables. We present the results in Table 3. Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, or reliable measurement of the covariate.

Table 3
Tests of Between-Subjects Effects. Dependent variable: achievement 2

Source	Type III Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
corrected model	26320.334[a]	7	5264.067	24.112	0.000	0.386
intercept	11226.116	1	11226.116	51.422	0.000	0.211
intervention	503.376	1	503.376	2.306	0.131	0.012
achievement 1	12409.039	1	12409.039	56.840	0.000	0.228
talk time 2	4.524	1	4.524	0.210	0.886	0.001
intervention* achievement 1	41.029	1	41.029	0.186	0.666	
intervention* talk time 2	2.819	1	2.819	0.013	0.910	
achievement 1* talk time 2	10.677	1	10.677	0.049	0.826	
intervention* talk time 2* achievement 1	4.894	1	2.447	0.011	0.989	0.000
error	41916.236	192	218.314			
total	711821.321	198				
corrected total	68236.570	197				

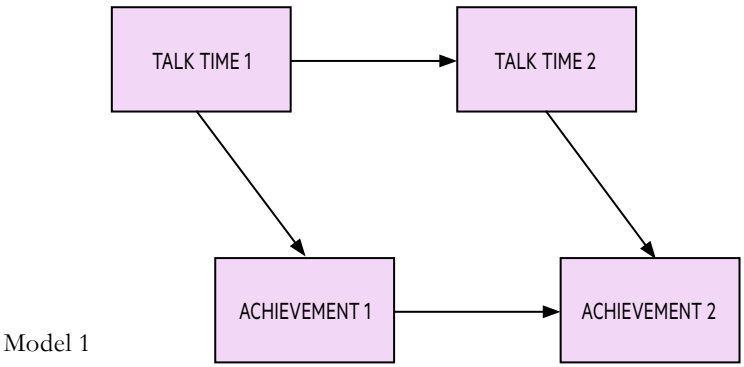
[a] *R* Squared = .386 (Adjusted *R* Squared = .370)

The ANOVA results indicate that intervention, as a predictor of "achievement 2" in reading literacy, loses its significance when controlling previous success rates in reading literacy ("achievement 1") and actual talk time ("talk time 2"). This is evidenced by the nonsignificant p values of the predictors and their interactions. The only statistically significant one is the initial performance of students, naturally. What does this suggest? Looking again at the descriptive indicators (Table 1), student talk time increased substantially in the intervention classes. There was no such change in the control classes. Achievement in reading literacy changed in both groups. Due to natural progression, students improved. In the intervention classes, the improvement is slightly higher. However, the improvement is less significant than the change in talk time. Thus, it seems that the changes in talk time did not have a definitive effect on achievement. What is the link between achievement and talk time?

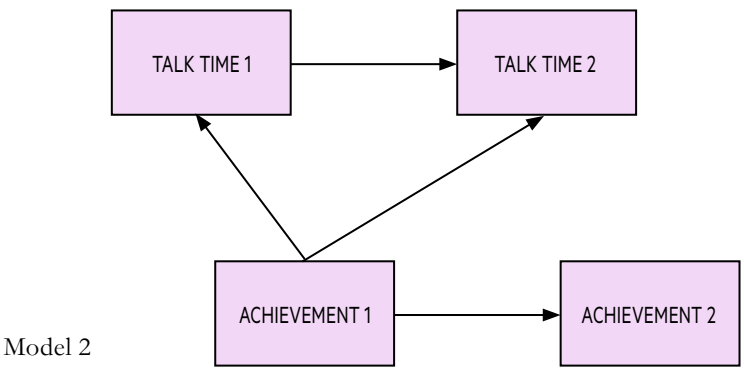
5.2 Is talk time a predictor of achievement, or is it the other way around?

The longitudinal nature of the data from our intervention study makes it possible to verify the multivariate relationships between student talk time and student achievement through path analyses. We assume that student talk time predicts student achievement and not vice versa. We test models defining different possibilities of relationships between variables. The aim is to identify which one best fits the data. In baseline Model 1 (see Figure 1), we assume that talk time in lessons of language arts ("talk time 1") positively predicts performance on reading literacy tests ("achievement 1"). Subsequently, "achievement 2" is explained through "talk time 2" and "achievement 1". However, the relationship between student talk time and student achievement may work in such a way that students with better achievement also talk in instructional communication more frequently. This is estimated in Model 2, where we reverse the direction of the association between "achievement 1" and "talk time 2" for this reason.

For the models, we use data from the full sample; that is, we combine intervention and control classes. Table 4 shows the basic characteristics of the variables entering the models.



Model 1



Model 2

Figure 1
Variants of models for path analysis

Table 4
Means and correlations among all variables selected for the path analysis

Variable	*N* (valid)	*M* (*SD*)	Correlation			
			1	2	3	4
1. talk time 1	239	12.4 (15.7)		0.24**	0.16*	0.14*
2. talk time 2	264	15.9 (17.1)	0.24**		0.10	0.05
3. achievement 1	221	49.9 (19.8)	0.16*	0.11		0.61**
4. achievement 2	237	56.4 (18.9)	0.14*	0.05	0.61**	

Note: Talk time – individual student talk time (average per one lesson in seconds); achievement – reading literacy (success in %). Pearson correlations: statistically significant at an alpha level *< 0.05; **< 0.01

It is clear from the mean values for both talk time and achievement variables that there were increases in both characteristics over the period. The reasons for the increases are explained in the previous question. Zero-order correlations show that the closest correlation is between "achievement 1" and "achievement 2". Other correlations, although not as strong, also confirm that testing the models outlined above through path analysis is meaningful.

Path analysis makes it possible to see the substantial direct and indirect effects of interactions between variables. At the same time, we can compare multiple options and decide which best fits the data based on their fit parameters. We used the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to compare models. We held to the basic rule that a model with lower AIC and BIC is more appropriate for the data (cf. Raftery, 1995). We offer the basic results of the two hypothetical models in Table 5.

Table 5
Fit indices of path analysis

							Our model		Saturated model	
Fit indices	χ^2	*df*	*p*	CFI	TLI	RMSE	AIC	BIC	AIC	BIC
Model 1	2.03	2	0.36	0.95	0.99	0.07	26.026	26.470	28.000	28.599
Model 2	7.472	3	0.05	0.91	0.86	0.08	29.472	29.879	28.000	28.519

Note: CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

We tested the direct path from talk time to student achievement in Model 1. The hypothesis that student achievement directly affects talk time was estimated in Model 2. According to the results of the chi-squared test and different structural equation modeling (SEM) criteria commonly used for SEM evaluation, Model 1 better fit the data. The statistical significance of the chi-squared test for Model 2 is crucial in this regard. The result indicates that we must reject the null hypothesis if the estimated model fits our data. For Model 1, on the other hand, we can maintain this hypothesis. The unsuitability of Model 2 was also confirmed by the information criteria (AIC and BIC). Here, for each of the models, the comparison of the estimated model (ours) with the so-called saturated model is essential. This is the model with the maximum number of parameters. Our model should always have lower values (the AIC and BIC values are always compared separately).

Model 2 does not meet this assumption. Model 1, on the other hand, shows good values of the other SEM criteria. The fit indices (AGFI, CFI, and TLI) were above the level of 0.95, considered a very good fit. On the other hand, the RMSEA measuring the misfit of the model only attained a threshold value of 0.08 (Kline, 2005).

Table 6
Unstandardized path coefficients, standard errors, and t-values for Model 1

Path	Estimate	*SE*	*t*	*p*
talk time 1 to achievement 1	0.209	0.086	2.432	<0.05
talk time 1 to talk time 2	0.283	0.068	4.180	<0.01
achievement 1 to achievement 2	0.598	0.051	11.811	<0.05
talk time 2 to achievement 2	0.005	0.059	0.081	0.93

Model 1 exhibits a relatively good fit to the data; however, it is on the border of acceptability. A detailed examination of the unstandardized coefficients and their statistical significance, as presented in Table 6, reveals that there is no significant relationship between "talk time 2" and "achievement 2". This means that the expected association between the repeated measure of talk time and the second reading literacy test does not exist within Model 1. Nevertheless, Model 1 does support the hypothesis that, in the absence of an intervention, "talk time 1" does indeed influence student "achievement 1". To refine our analysis, we adjusted Model 1. This involved removing an ineffective link and introducing a direct connection between "talk time 1" and "achievement 2". Talk time 1 is considered a long-lasting characteristic and is expected to directly impact student achievement, even with a longer time lag.

Table 7
Fit indices of path analysis (Model 3)

							Our model		Saturated model	
Fit indices	χ^2	*df*	*p*	CFI	TLI	RMSE	AIC	BIC	AIC	BIC
Model 3	0.637	2	0.73	0.99	1.06	0.00	24.63	25.08	29.000	28.519

Note: CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike's information criterion; BIC = Bayesian information criterion.

We summarize the results of this final model in Table 6. We described the SEM criteria showing a good fit to the data. From the results presented in Table 6, Model 3 indicates an excellent fit. Compared to Model 1, the change in the RMSE criterion (0.00) is particularly significant, marking a low misfit in the final Model 3. Therefore, we consider Model 3 as the final model. The correlations and their strength as represented by the standardized beta coefficients are presented in Figure 2.

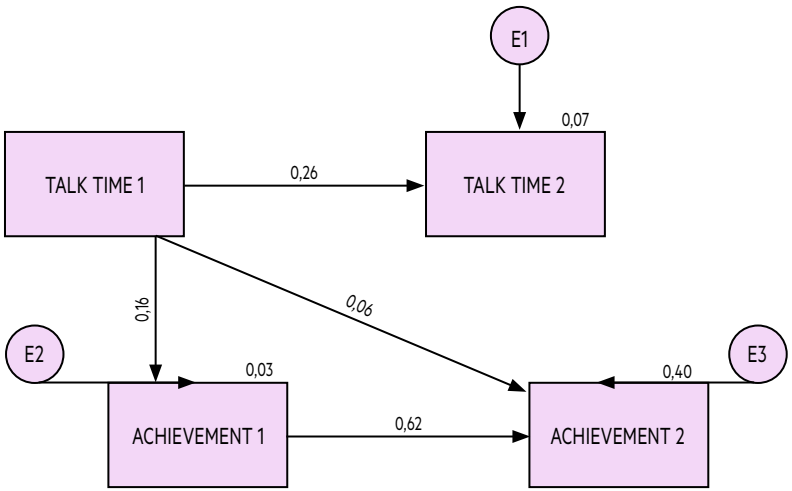


Figure 2
Results for final model relating participation in classroom discourse and student achievement with standardized coefficients

Source: IBM SPSS AMOS

In the final model, the outcome variable ('achievement 2') is explained by all included factors, with an Adjusted R^2 of approximately 40%. This is a perfect result. The other endogenous variables of the model do not reach such values. The first round of student achievement measurement is explained at 3%. The resulting R^2 for the second measurement of participation in classroom discourse ("talk time 2") reached a value of 7%. In both cases, this was due to only one predictor: the initial characteristic of participation in classroom discourse ("talk time 1"). The more detailed analysis allows us to analyze the direct, indirect, and total causal effects of variables. We summarize the results in Table 8.

Table 8
Direct, indirect, and total effects (standardized) for the path model

Path	Direct effect	Indirect effect	Total
Talk time 1 to achievement 1	0.205	0.000	0.205
Talk time 1 to talk time 2	0.281	0.000	0.281
Achievement 1 to achievement 2	0.588	0.000	0.588
Talk time 1 (via achievement 1) to achievement 2	0.077	0.163	0.240

The most critical finding confirms the direct positive effect of talk time on student achievement. Higher talk time means improving achievement in reading literacy. In the first round of measurement, the strength is expressed by a standardized regression coefficient of 0.16 (the unstandardized regression coefficient has a value of 0.21, which, given the units of measure, can be interpreted as meaning that an increase of 5 seconds means a 1% better success rate in the test). The confirmed relationships between the first and second rounds of measurement for both student talk times and achievements are logical, given the pairing of the measure.

The second significant finding does not confirm the connection between "talk time 2" and "achievement 2". This lack of association can be attributed to the relatively short time between measurements. It appears that changes in talk time did not have a significant impact on achievement within this timeframe. It is worth noting educational changes typically exhibit longitudinal patterns (Larraín et al., 2018); this is supported by our observation of the relationship between "talk time 1" and "achievement 2". The initial characteristic of student "talk time 1" has a lasting influence on "achievement 1," demonstrating a strong connection. Although the direct influence on "achievement 2" is relatively weak (0.06), this link is bolstered indirectly through its effect on "achievement 1" (0.163). Consequently, the overall impact of "talk time 1" on "achievement 2" remains substantial and statistically significant (0.240). These findings strongly suggest that changes in "talk time 2" are likely to manifest in future academic achievements, specifically "achievement 3".

6 Discussion

The present study is the first to investigate the link between student participation in classroom discourse and student achievement within the frame of an intervention program focused on increasing and equalizing

student participation. Although several intervention projects have been conducted to influence student achievement by changing the quality of classroom discourse (Alexander, 2018; Hardman, 2016; Howe et al., 2019; Ruthven et al., 2017) none of them has yet taken into account individual differences among students.

There is some evidence indicating a positive relationship between individual participation in classroom discourse and student achievement (Decristan et al., 2023; Ing et al., 2015; Neuman et al., 2021; Rüede et al., 2023; Schnitzler et al., 2020; Šeďová et al., 2019; Webb et al., 2014); however, none of the previous studies was designed to control the causality in this link. The interventional nature of our data allowed us to test two models, one expecting achievement to be affected by talk time and the second expecting talk time to be affected by achievement. Simply said, we asked whether students perform better due to their extensive talk or talk more due to their good achievement. Our analysis confirmed the first model to better fit the data: the quantity of student verbal participation predicts their achievement.

The effect size of talk time is quite small, but it is a piece in the mosaic of other findings confirming the predictors of student achievement (Hattie, 2009; King et al., 2024; Mullis et al., 2001; Terhart, 2011). Determining the predictors of achievement is essential in education as it makes it possible to think about how to create conditions for students that will increase their chances of success. Some of the recognized predictors cannot be influenced – such as socioeconomic background, previous schooling, and prior achievement. Others can be – including learning motivation and academic self-concept – but it is a complex and challenging task. Student participation in classroom discourse is prone to change in stimulative conditions, as evidenced by our data. Therefore, it deserves the careful attention of educators.

Our findings imply that teachers should invite all students to participate in classroom discourse for the sake of their learning. The goal is to equalize the uneven participation opportunities for different students in the same class (Vrikki et al., 2019). It has been established that students who are both silent and disengaged face learning challenges (Bae & DeBusk-Lane, 2019; Schnitzler et al., 2020). Therefore, a key task for teachers is to assist these students in finding their voices, thereby enhancing their learning opportunities. This becomes particularly relevant in the Czech educational culture, where students are expected to be attentive but not necessarily outspoken (see Šeďová & Sedláček, 2023). To ensure effective learning for all students, it is essential to challenge and move beyond these traditional norms.

Our analysis showed that students who had participated in an intervention enhanced their performance in literacy tests more than students in control classes. However, the change in achievement was not adequately large

compared to the change in participation. This difference raises the question of what this result means considering the finding that talk time predicts achievement. We hypothesize that the change in verbal participation needs more time to be reflected in student achievement. This hypothesis may be supported by the fact that achievement in post-measurement was more affected by the talk time before the intervention than at the end of the intervention.

Student participation has been found to be structured into quite stable patterns co-created by students, teachers, and peers in the classroom (Kovalainen & Kumpulainen, 2007; Šeďová & Sedláček, 2023). The influence of talk on achievement is thus long standing. It can be assumed that talk does not have an immediate effect, but rather a long-term cumulative impact. The intervention we conducted in participating classrooms led to a reconstruction of participation patterns with a strong participation elevation in the previously silent students and a mild decrease in the previously most vocal students. Several future scenarios could be relevant. First, the new participation patterns become stable and after some time the increased participation transforms into improved performance. Second, new participation patterns will evaporate when not supported by the intervention team and therefore there will be no change in achievement. Third, the new participation patterns will survive due to teacher commitment to equitable participation, but they will not affect student learning.

These three possible scenarios outline the agenda for future research. It is important to continue investigating the potential influences of student participation in classroom discourse, as recent findings are promising and indicate easy-to-implement tools for enhancing student learning. When designing future studies, it is necessary to include long-term monitoring of both student participation and student achievement. Only through delayed post-measurements can we find whether promoting verbal participation could be conducive to changes in their achievement.

Acknowledgments

This article is an output of the project *Collectivity in Dialogic Teaching: An Intervention Study* (GA21-16021S) funded by the Czech Science Foundation. This output was supported by the NPO "Systemic Risk Institute" number LX22NPO5101, funded by the European Union – Next Generation EU (Ministry of Education, Youth and Sports of the Czech Republic, NPO: EXCELES).

References

Alexander, R. (2018). Developing dialogue: Genesis, process, trial. *Research Papers in Education, 33*(5), 561–598.
https://doi.org/10.1080/02671522.2018.1481140

Bae, L. Ch., & DeBusk-Lane, M. (2019). Middle school engagement profiles: Implications for motivation and achievement in science. *Learning and Individual Differences, 74*(2), 101753.
https://doi.org/10.1016/j.lindif.2019.101753.

Böheim, R., Urdan, T., Knogler, M., & Seidel, T. (2020). Student hand-raising as an indicator of behavioral engagement and its role in classroom learning. *Contemporary Educational Psychology, 62*, 101894.
https://doi.org/10.1016/j.cedpsych.2020.101894

Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*(5), 365–374.
https://doi.org/10.1037/h0029908

Caspi, A., Chajut, E., Saporta, K., & Beyth-Marom, R. (2006). The influence of personality on social participation in learning environments. *Learning and Individual Differences*, *16*(2).
https://doi.org/10.1016/j.lindif.2005.07.003

Clarke, S. N. (2015). The right to speak. In L. B. Resnick, C. S. C. Asterhan & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 167–180). American Educational Research Association.

Chang, D. F., Chien, W. C., & Chou W. C. (2016). Meta-analysis approach to detect the effect of student engagement on academic achievement. *ICIC Express Letters, 10*, 2441–2446.
https://doi.org/10.24507/icicel.10.10.2441.

Decristan, J., Jansen, N. C., & Fauth, B. (2023). Student participation in whole-class discourse: individual conditions and consequences for student learning in primary and secondary school. *Learning and Instruction, 86*, 101748.
https://doi.org/10.1016/j.learninstruc.2023.101748

Erath, K., Prediger, S., Quasthoff, U., & Heller, V. (2018). Discourse competence as important part of academic language proficiency in mathematics classrooms: the case of explaining to learn and learning to explain. *Educational Studies in Mathematics, 99*, 161–179.
https://doi.org/10.1007/s10649-018-9830-7

Hardman, J. (2016). Opening-up classroom discourse to promote and enhance active, collaborative and cognitively-engaging student learning experiences. In C. Goria, O. Speicher & S. Stollhans (Eds.*), Innovative language teaching and learning at university: enhancing participation and collaboration* (pp. 5–16).
https://doi.org/10.14705/rpnet.2016.9781908416322

Hattie, J. (2009). *Visible Learning. A Synthesis of Over 800 Meta-Analysis Relating to Achievement.* Routledge.
https://doi.org/10.4324/9780203887332

Helgevold, N. (2016). Teaching as creating space for participation – establishing a learning community in diverse classrooms. *Teachers and Teaching, 22*(3), 315–328.

Hennessy, S-, Calcagni, E., Leung, A., & Mercer, N. (2023). An analysis of the forms of teacher-student dialogue that are most productive for learning. *Language and Education, 37*(2), 186–211. https://doi.org/10.1080/09500782.2021.1956943

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher–student dialogue during classroom teaching: Does it really impact upon student outcomes? *Journal of the Learning Sciences, 28*(4-5), 462–512. https://doi.org/10.1080/10508406.2019.1573730

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Inagaki, K., Hatano, G., & Morita, E. (1998). Construction of mathematical knowledge through whole-class discussion. *Learning and Instruction*, *8*(6). https://doi.org/10.1016/S0959-4752(98)00032-2

Ing, M., Webb, N. M., Franke, M. L., Turrou, A. C., Wong, J., Shin, N., & Fernandez, C. H. (2015). Student participation in elementary mathematics classrooms: the missing link between teacher practices and student achievement? *Educational Studies in Mathematics, 90*, 341–356. https://doi.org/10.1007/s10649-015-9625-z

Jurik, V., Gröschner, A., & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. *Learning and Instruction, 23*, 33–42.

Kawabe, R., Yamamoto, M., Aoyagi, S., & Watanabe, T. (2014). Measurement of hand raising actions to support students' active participation in class. In S. Yamamoto (Ed.), *Human Interface and the Management of Information. Information and Knowledge Design and Evaluation*. Lecture Notes in Computer Science (Vol. 8521). Springer. https://doi.org/10.1007/978-3-319-07731-4_20

Kelly, S. (2008). Race, social class, and student engagement in middle school English classrooms. *Social Science Research*, *37*(2). https://doi.org/10.1016/j.ssresearch.2007.08.003

King, R. B., Cai, Y., & Elliot, A. J. (2024). Income inequality is associated with heightened test anxiety and lower academic achievement: A cross-national study in 51 countries. *Learning and Instruction*, *89*. https://doi.org/10.1016/j.learninstruc.2023.101825

Kline, R. B. (2005). *Principles and practice of structural equation modelling.* Guilford Press.

Kovalainen, M., & Kumpulainen, K. (2007). The social construction of participation in an elementary classroom community. *International Journal of Educational Research, 46*(3–4), 141–158. https://doi.org/10.1016/j.ijer.2007.09.011

Kyriakides, L., Creemers, B., Panayiotou, A., & Charalambous, E. (2020). *Quality and equity in education: Revisiting theory and research on educational effectiveness and improvement*. Routledge.

Larrain, A., Freire, P., Grau, V., López, P., Salvat, I., Silva, M., & Gastellu, V. (2018). The effect of peer-group argumentative dialogue on delayed gains in scientific content knowledge. *New Directions for Child and Adolescent Development*, *162*, 67–87. https://doi.org/10.1002/cad.20263

Morlà-Folch, T., Renta Davis, A., I., Padrós, M., & Valls, R. (2022). A research synthesis of the impacts of successful educational actions on student outcomes. *Educational Research Review, 37*(1). https://doi.org/10.1016/j.edurev.2022.100482

Moser, M., Zimmermann, M., Pauli, C., Reusser, K., & Wischgoll, A. (2022). Student's vocal participation trajectories in whole-class discussions during teacher professional development. *Learning, Culture and Social Interaction, 34*, 100633. https://doi.org/10.1016/j.lcsi.2022.100633

Mullis, I. V. S., Martin, M. O., & Gonzales, E. J. (2001). *International Achievement in the Processes of Reading Comprehension: Results from PIRLS 2001 in 35 countries.*

Mundelsee, L., & Jurkowski, S. (2021). Think and pair before share: Effects of collaboration on students' in-class participation. *Learning and Individual Differences, 88*, 102015. https://doi.org/10.1016/j.lindif.2021.102015

Myhill, D. (2002). Bad boys and good girls? Patterns of interaction and response in whole class teaching. *British Educational Research Journal*, *28*(3). https://doi.org/10.1080/01411920220137430

Neuman, S. B., Samudra, P., & Danielson, K. (2021). Effectiveness of scaling up a vocabulary intervention for low-income children, pre-K through first grade. *Elementary School Journal, 121*(3), 385–409. https://doi.org/10.1086/712492

O'Connor, C., Michaels, S., Chapin, S., & Harbaugh, A. G. (2017). The silent and the vocal: Participation and learning in whole-class discussion. *Learning and Instruction, 48*, 5–13. https://doi.org/10.1016/j.learninstruc.2016.11.003

Orner, A., & Netz, H. (2023). Taking, begging, or waiting for the floor: students' social backgrounds, entitlement and agency in classroom discourse. *Discourse: Studies in the Cultural Politics of Education, 44*(2), 221–237. https://doi.org/10.1080/01596306.2021.1989573

Panayiotou, A., Herbert, B., Sammons, P., & Kyriakides, L. (2021). Conceptualizing and exploring the quality of teaching using generic frameworks: A way forward. *Studies in Educational Evaluation*, *70*, 1–15. https://doi.org/10.1016/j.stueduc.2021.101028

Rüede, C., Streit, C., Mok, S. Y., & Laubscher, R. (2023). Orchestrating productive classroom talk in Swiss second grade mathematics classrooms. *Journal für Mathematik-Didaktik, 44*, 385–415. https://doi.org/10.1007/s13138-023-00224-2

Ruthven, K., Mercer, N., Taber, K. S., Guardia, P., Hofmann, R., Ilie, S., Luthman, S., & Riga, F. (2017). A research-informed dialogic-teaching approach to early secondary school mathematics and science: the pedagogical design and field trial of the epiSTEMe intervention. *Research Papers in Education, 32*(1), 18–40. https://doi.org/10.1080/02671522.2015.1129642

Schenke, K. (2018). From structure to process: Do students' own construction of their classroom drive their learning? *Learning and Individual Differences, 62*, 36–48. https://doi.org/10.1016/j.lindif.2018.01.006.

Sedláček, M. & Šeďová, K. (2020). Are student engagement and peer relationships connected to student participation in classroom talk? *Learning, Culture and Social Interaction*, 26. https://doi.org/10.1016/j.lcsi.2020.100411

Šeďová, K., & Sedláček, M. (2023). How vocal and silent forms of participation in combination relate to student achievement. *Instructional Science*, *51*, 343–361. https://doi.org/10.1007/s11251-022-09609-1

Šedová, K., Sedláček, M., Švaříček, R., Majcík, M., Navrátilová, J., Drexlerová, A., Kychler, J., & Šalamounová, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning and Instruction*, *63*. https://doi.org/10.1016/j.learninstruc.2019.101217

Shepherd, A. M. (2014). The discursive construction of knowledge and kvity in classroom interactions. *Linguistics and Education, 28*, 79–91. https://doi.org/10.1016/j.linged.2014.08.006.

Schnitzler, K., Holzberger, D., & Seidel, T. (2020). All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education, 36*, 627–652. https://doi.org/10.1007/s10212-020-00500-6

Terhart, E. (2011). Has John Hattie really found the holy grail of research on teaching? An extended review of visible learning. *Journal of Curriculum Studies*, *43*(3), 425–438. https://doi.org/10.1080/00220272.2011.576774

Young, C. (2014). Predictors of quality verbal engagement in third-grade literature discussions. *International Electronic Journal of Elementary Education*, *6*(3).

Vrikki, M., Kershner, R., Calcagni, E., Hennessy, S., Lee, L., Hernández, F., Estrada, N., & Ahmed, F. (2019). The teacher scheme for educational dialogue analysis (T-SEDA): developing a research-based observation tool for supporting teacher inquiry into pupils' participation in classroom dialogue. *International Journal of Research & Method in Education, 42*(2), 185–203. https://doi.org/10.1080/1743727X.2018.1467890

Webb, N. M., Franke, M. L., Ing, M., Wong, J., Fernandez, C. H., Shin, N., & Turrou, A. C. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research*, *63*. https://doi.org/10.1016/j.ijer.2013.02.001

STUDIA PAEDAGOGICA

vol. 29, n. 3, 2024
https://doi.org/10.5817/SP2024-3-3

STUDY

FROM POLICY TO PRACTICE: MONITORING SDG TARGETS WITH PIRLS 2021 AUSTRIAN DATA

Surette van Staden[a], Sunet Grobler[a]

[a] Institute for Teacher Education and School Research, Faculty of Teacher Education, University of Innsbruck

ABSTRACT

This study uses the Progress in International Reading Literacy Study (PIRLS) 2021 data to test quality education targets related to primary education achievement, early childhood, skills for work, equity, the learning environment, and teachers' highest levels of qualification. Using Austrian data as a specific case, a regression analysis was performed on items from the student, home, and teacher questionnaires with overall reading literacy achievement as the outcome variable. Results show a drop in Austrian primary education reading literacy achievement between PIRLS 2016 and PIRLS 2021. Targets of early childhood, skills for work, equity, and a safe and effective learning environment are statistically significantly related to reading achievement. Boys trail girls in reading achievement, and higher teacher qualifications do not translate into significantly better reading results. The study's focus on targets with evidence from PIRLS raises two issues: the extent of the PIRLS contribution to measuring and monitoring SDG 4 targets and how PIRLS can help operationalize interconnected targets.

KEYWORDS

International Large-Scale Assessments (ILSAs); Progress in International Reading Literacy Study (PIRLS); Sustainable Development Goals (SDGs); Sustainable Development Goal 4 (SDG 4)

CORRESPONDING AUTHOR
Surette van Staden, Faculty of Teacher Education, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria
e-mail: surette.van-staden@uibk.ac.at

Introduction

The increasing globalization of education standards and market demands require more focus on the quality of education. According to Tatto and Pippin (2017), it has become a significant and fiercely debated area that cannot be disregarded. "Education is high on the agenda of governments around the globe" (Robinson, 2016, p. 6) as nations move to upgrade teachers and reform teaching to improve their standings on international assessments. The global pursuit of the Sustainable Development Goals (SDGs), led by agencies like the United Nations Development Program (UNDP), aims to reach the SDGs by 2030 and highlights SDG 4 as a pivotal driver in achieving this goal. The 17 SDGs represent goals for social improvement (e.g., hunger, poverty, health, well-being, and reduced inequalities), the environment (e.g., climate action, clean energy, life below water, and life on land), and economic development (e.g., industry, innovation and infrastructure, decent work, and economic growth) (UN, 2015). To gauge evidence of the achievement of the SDGs in the European Union (EU), a systematic literature review conducted by Trane, Marelli, Siragusa, Pollo, and Lombardi (2023) highlighted the rising interest of scholars in operationalizing Agenda 2030. European studies currently devote major interest to environmental concerns (especially linked to SDG 6, 7, 12, 13, and 15), while social issues (e.g., SDG 4, 5, and 10) still warrant more research. While the EU strongly committed itself to the SDGs, clear metrics and data are essential for countries to track progress and achieve these goals. Trane et al. (2023) mentioned, as examples of what has been done in the EU thus far, the "European Sustainable Development Report" (SDSN & IEEP, 2021), which has been monitoring the performance of all EU members, the United Kingdom, partner countries, and the EU as a whole since 2019. The OECD published the "Measuring Distance to SDG Targets" report in 2016, 2017, 2019, and 2022 (OECD, 2022), grouping national trends toward the SDGs. The "Monitoring Report on Progress Towards the SDGs in an EU Context" is published yearly (Eurostat, 2022), with analysis that builds on the EU SDG indicator set, 100 indicators developed in cooperation with a large number of stakeholders for the specific EU context and structured along the 17 SDGs. Finally, "Measuring the Situation of the European Union with regard to the SDGs" (ASviS, 2019) by the Italian Alliance for Sustainable Development tracks the progress of the EU on each SDG by a subset of Eurostat indicators, covering the period from 2010 as a baseline up until 2017.

SDG 4 is crucial in achieving the remaining SDGs (Madalinska-Michalak, 2023; OECD, 2017; Priyadarshini, 2019; UNESCO, 2021; UN, 2018). It significantly empowers individuals to develop knowledge, skills, and values

that promote the SDGs. This idea is supported by the fact that education is recognized as a crucial tool for achieving the SDGs and improving people's capacities to address environmental and development issues. The global community has mandated this recognition, and the UN has emphasized the importance of education since 1992. Several studies, including those conducted by Fehlner (2019), Havea and Mohanty (2020), Kumar (2020), and the OECD (2017), have highlighted the critical role of education in promoting sustainable development. SDG 4 represents quality education, a goal that broadens opportunities across all phases, including primary, secondary education, vocational, higher, and adult education to encompass outcomes of literacy, numeracy, and wider aspirations such as citizenship, sustainability, and gender equality (Bruns et al., 2019; Unterhalter, 2019). This goal was endorsed by Priyadarshini (2019), who stated that education, literacy, and adult learning are key to achieving the SDGs. In summary, SDG 4 is crucial in promoting sustainable development, empowering individuals, and achieving the remaining SDGs by providing quality education.

The concept of quality education (which is also the label for SDG 4) is too complex and multifaceted to define, especially in the field of teacher education research. This view was also held by Flores (2023, p. 32), who stated that "there is no single definition of quality that applies universally nor is there a single recipe for improving quality in initial teacher education programmes." Therefore, this research seeks to explore and define the dimensions of quality education within the context of SDG 4 by investigating the relationships between early home literacy activities, the learning environment, equity considerations, and reading literacy outcomes among ISCED level 2 (i.e. fourth grade) Austrian students, utilizing data from the Progress in International Reading Literacy Study (PIRLS) 2021. For purposes of the current analyses, Austria presents itself as a case of a developed, central European country with a history of PIRLS participation since its 2001 cycle of administration. World Bank figures show that Austria has a total population of 8.9 million inhabitants (World Bank, 2021, as cited in Van Staden & Schreiner, 2023), with an overall student population of approximately 1.1 million children in the 2020/2021 school year (Statistik Austria, 2022). The Gross Domestic Product (GDP) for 2021 is listed as 477.08 billion US dollars in total and 53,267.9 US dollars per capita (World Bank, 2021, as cited in Van Staden & Schreiner, 2023). Depending on the educational track, the government provides up to 13 years of schooling and one additional year of compulsory kindergarten. By examining these factors, this research aims to contribute to a nuanced understanding of how to conceptualize and operationalize quality education, particularly in terms of fostering reading literacy skills and addressing disparities in educational achievement.

First, this paper introduces the SDG targets and their indicators, followed by the applicability of ILSAs and their relation to SDG 4. It then discusses the research questions and methodology, followed by the results, discussion, limitations, and conclusion of the study.

1 SDG 4 targets and indicators

SDG 4 aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (UN, 2015, p. 19). This SDG consists of 10 targets and 12 indicators as suggested by the UN (2015). The first seven targets with indicators are based on the outcomes that are envisioned by the targets, and the last three focus on the means of implementation. Outcome Target 4.1 and Target 4.2 aim to provide unrestricted access to quality education that prepares students for their future education and career paths. While Indicator 4.2.2 includes participation rates, data on these rates are not available in PIRLS. Therefore, the current research does not provide further insight into this indicator. Target 4.3 aims to achieve gender equality through empowering all women and girls, based on equal technical, vocational, and tertiary education access. Target 4.4 focuses on student readiness for the professional world and educational access for marginalized groups. Target 4.5 strives to provide equal access to education for all individuals, regardless of any form of discrimination they may face. This is particularly crucial for those who have been historically marginalized and may encounter further obstacles to education. The objective of Target 4.5 is to promote equal educational opportunities and foster a more equitable society in which everyone can achieve their aspirations. Target 4.6 aims to elevate literacy and numeracy levels across all age groups, including adults and youth. Finally, Target 4.7 emphasizes education on certain content, knowledge, and skills that contribute to sustainable development, human rights, gender equality, and cultures of peace and non-violence.

The last three targets provide the means for executing the quality and equality targets of the first seven (Sayed & Moriarty, 2020). Target 4.a aims at creating effective and inclusive learning environments that are safe and gender sensitive. This target can be achieved by building and upgrading education facilities for children and people with disabilities. Target 4.b aims to increase scholarships for vocational training in information and communications technology (ICT), technical, engineering, and scientific programs. Target 4.c, as an implementation target, aims to increase the number of qualified teachers by supporting underdeveloped countries through international cooperation and other means. Since these targets, together with

their indicators, guide stakeholders in achievement and progress (Moldan & Dahl, 2007), they can also be linked to ILSAs and the achievement of quality education in schools.

2 The role of ILSAs in quality education

The International Association for the Evaluation of Educational Achievement (IEA) develops and conducts ILSAs globally to show student achievement in education systems (Mullis et al., 2023; Leino et al., 2022). ILSAs include IEA studies such as PIRLS (Progress in International Reading Literacy Study), TIMSS (Trends in International Mathematics and Science Study), and the Organization for Economic Cooperation and Development's (OECD) PISA (Program for International Student Assessment) and TALIS (Teaching and Learning International Survey) studies. International assessments are instrumental in monitoring education policies and practices by providing comparative data on educational outcomes and practices across countries; they have been described as the "global yardstick for measuring success in education" (Schleicher, 2017, p. 123 in Ledger et al., 2019). These assessments often inform discussions and decisions regarding education reform and improvement efforts (Mullis et al., 2023). ILSAs are crucial in providing valuable insights into the education systems across countries, facilitating discussions, and shaping policy decisions to improve educational outcomes.

According to Robinson (2016), education is a top priority for governments worldwide, and ILSAs provide valuable evidence to support, monitor, and benchmark educational development. Addey and Sellar (2019) explained that although there was initial skepticism about ILSAs, they have become an essential tool for policymaking over the last two decades. Governments are willing to invest a lot of money in ILSAs because they provide reliable data to evaluate educational outcomes and identify effective policies. Participating in ILSAs demonstrates that a country shares common educational values and goals with other participating nations. Addey and Sellar (2019) developed a four-dimensional framework that outlines the reasons for government participation, including political, economic, technical, and sociocultural rationales.

ILSAs have several rationales that governments utilize for their strategic benefits (Addey & Sellar, 2019). Political rationales involve the use of ILSAs as a tool to navigate domestic policy landscapes, reconcile policy disputes, gain political support, and differentiate political agendas (Addey & Sellar, 2019). The outcomes of ILSAs are often leveraged to instigate policy reforms and shape public narratives around education (Addey & Sellar, 2019;

Waldow, 2017). Economic rationales recognize the link between skill measurements by ILSAs and economic growth, making ILSA results a crucial element for economic strategies (Addey & Sellar, 2019). Technical rationales refer to the methodological rigor and capacity-building potential of ILSAs in educational assessments (Addey & Sellar, 2019). This enables countries to enhance their technical expertise in developing, implementing, and analyzing comprehensive learning evaluations (Addey & Sellar, 2019). Lastly, sociocultural rationales involve the use of ILSAs to align countries with international norms and models for modern statehood. This displays a commitment to modern education systems and policies (Addey & Sellar, 2019).

The use of ILSAs by governments offers a multi-dimensional approach to education policy-making that encompasses political, economic, technical, and sociocultural perspectives. One such ILSA includes the IEA's PIRLS, a global initiative to improve reading, teaching, and learning. Since its establishment in the early 2000s, PIRLS has been administered every five years to assess children's reading comprehension following four years of formal education (Mullis, et al., 2012).

PIRLS assesses the reading literacy of fourth-grade students using two main types of reading tasks and four comprehension methods that evaluate student reading ability in traditional and online formats (Mullis & Martin, 2019). The evaluation focuses on each type of reading task and comprehension method, breaking down what portion of the test is dedicated to each aspect (Mullis & Martin, 2019). The purposes for reading include literary experience and acquiring and using information; the processes include focusing on and retrieving explicitly stated information, making straightforward inferences, interpreting and integrating ideas and information, and evaluating and analyzing content and textual elements (Mullis & Martin, 2019). PIRLS thoroughly and comprehensively assesses student reading literacy and reflects the complexity of reading skills which are essential everywhere.

3 The systemic interconnectedness of PIRLS in relation to SDG 4 targets

The PIRLS 2021 assessment framework recognizes the impact of both student achievement and contextual background factors on the learning environment (Mullis & Martin, 2019). Figure 1 presents the PIRLS 2021 assessment framework. It acknowledges the interplay between student achievement (as an important outcome of quality education) and contextual background factors that shape the learning environment. Figure 1 illustrates this interconnectedness as follows:

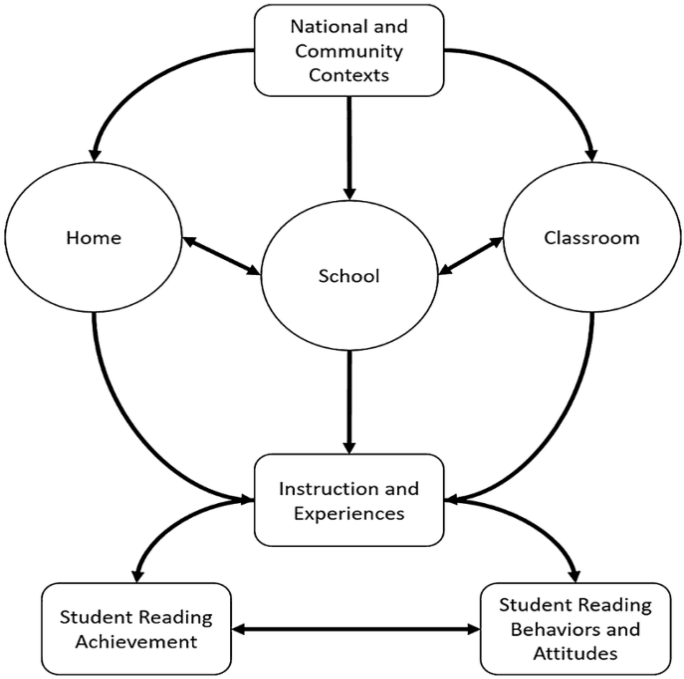


Figure 1
PIRLS 2021 assessment framework Mullis et al. (2012, p. 35)

The assessment framework in Figure 1 further illustrates the interconnected nature of student reading achievement (literacy) as well as their reading behaviors and attitudes towards reading, considering different contexts in which reading takes place, including the home, school, and classroom within national and community contexts. While this assessment framework refers to the work of Mullis et al., dating back to 2012, the tenets and design of this framework remain the same for PIRLS studies in consequent cycles in 2016 and 2021. The framework depicted above, illustrating the complex relationship between student reading achievements and their attitudes towards reading across various contexts, serves as a compelling entry point to discuss the broader interconnectedness of SDG 4 targets and the SDGs, emphasizing the foundational role of literacy as an outcome as one dimension of achieving comprehensive educational and developmental outcomes (Priyadarshini, 2019). Figure 2 illustrates how SDGs are linked to one another.

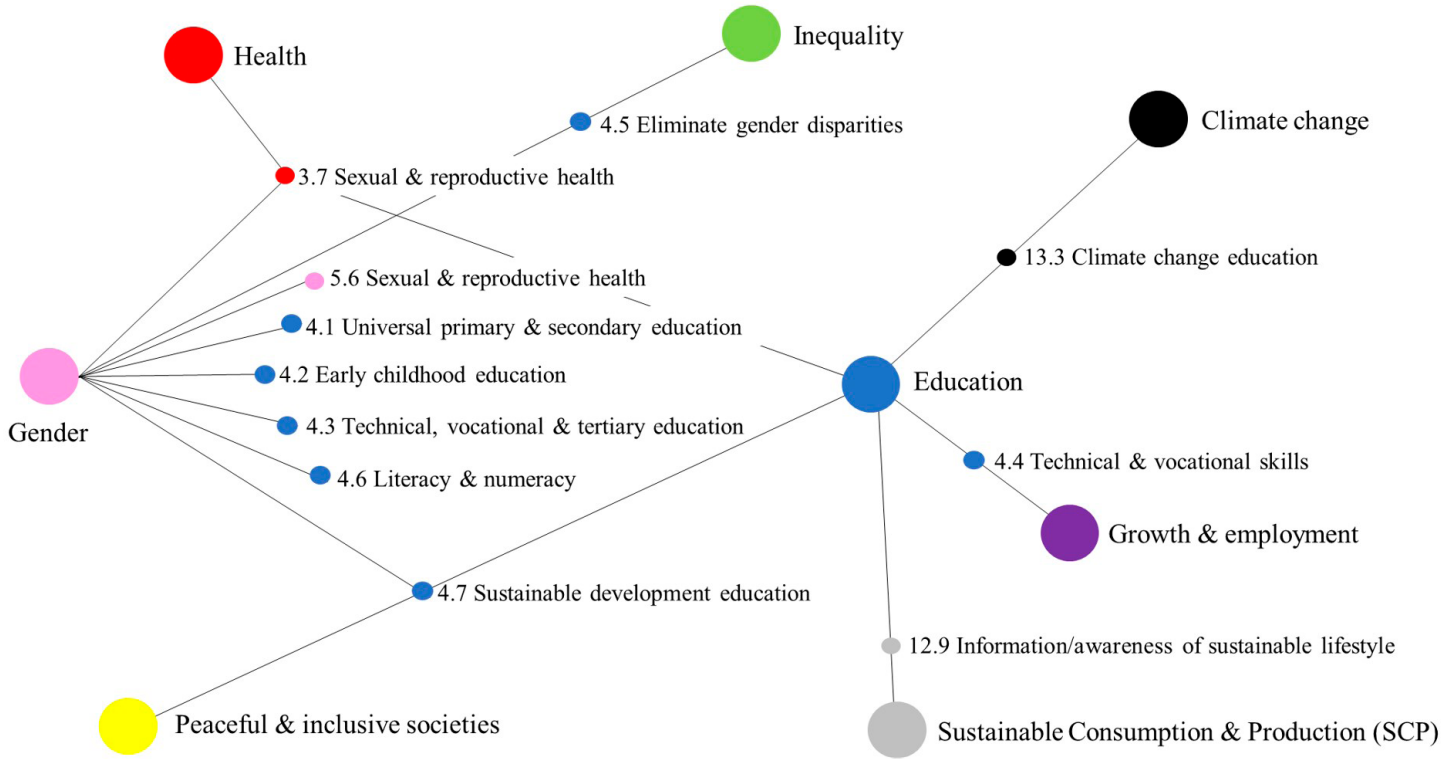


Figure 2

The links between education and other SDG targets

Thaung (2018, p. 17)

Figure 2 shows the centrality of education by linking it to different SDGs. The figure shows the links between education, gender, and overall equality, based on Targets 4.1, 4.2, 4.3, 4.5, 4.6, and 4.7 which are present in the quality education SDG but also concerns equality which links it to the other equality SDGs such as SDGs 5, 8, 10, 11, and 16. Target 4.4 links education with growth and employment. This diagram is selected to show how interconnected the SDGs are and the significance of the role SDG 4 plays in Agenda 2030 (also according to Grobler & Dittrich, 2024; Madalinska-Michalak, 2023; OECD, 2017; Thangeda et al., 2016; UNESCO, 2023; UN, 2018). Thaung (2018) highlighted the deep links between education (central to SDG 4) and other key areas like health (SDG 3), climate action (SDG 13), and inclusive and peaceful societies (SDG 16). These interconnections underscore the multifaceted role of education in promoting health, gender equality, environmental awareness, economic skills, and inclusive and peaceful societies, further emphasizing education's pivotal role in achieving the SDGs (Thaung, 2018).

Education in itself has many outcomes, one of which is literacy. Priyadarshini (2019) supports the idea that literacy has a positive impact on both social and economic aspects of an individual life and plays a significant role in the development of communities and nations. Insufficient literacy can significantly hinder an individual's involvement in the lifelong learning process, which is crucial for their growth and development (Priyadarshini, 2019). Literacy is not viewed as an enabler but rather an outcome in this study, since the SDG 4 targets can be viewed as essential preconditions for literacy. For instance, early childhood development (Target 4.2) is widely recognized as foundational for later literacy, as early cognitive and socio-emotional support improves children's readiness for learning to read and write (UNESCO, 2015).

Teachers play a crucial role in providing quality education, which makes Target 4.c important in supporting the other SDGs and SDG 4 targets. According to UNESCO (2017), SDG 4 is directly linked to SDG Targets 3.7, 5.6, 8.6, 12.8, and 13.3 through education relations. UNESCO (2017) has also emphasized the urgency of Target 4.c, stating that "teachers are the key to achieving all the SDG targets" (p. 15). This target is essential in emphasizing the importance of quality education in achieving the SDGs. Given its direct impact on achieving other SDGs and SDG 4 targets, the significance of Target 4.c cannot be overstated. Therefore, it is essential to address the urgent need to increase the supply of qualified teachers to ensure quality education for everyone.

The targets and indicators guide the progress and achievement of SDG 4 worldwide. The indicators are useful tools to manage important dimensions of the environment and society (Dahl, 2012). These indicators can be related to early home literacy activities, the learning environment, equity considerations,

and their impact on reading literacy outcomes, tying them back to the specific targets and indicators of SDG 4. Quality education as a goal comprises several targets; in operationalizing these targets, PIRLS can be instrumental in tracking and monitoring performance and mobilizing the associated metrics that better depict the key tenets of the goal (IEA, 2021). These include:

3.1 Primary education (Target 4.1)

By employing international benchmarks, PIRLS can provide diagnostic evidence of children's reading comprehension skills and abilities when they reach the fourth year of primary school, which translates to fourth grade in most countries (IEA, 2016). PIRLS uses four benchmarks: achievement at the Low International Benchmark, where students are only able to achieve at or below 400 score points; the Intermediate International Benchmark, with achievement at 475 score points; the High International Benchmark, with achievement at 550 score points, and the Advanced International Benchmark, with achievement at or above 625 score points (Mullis et al., 2017). In this study, the PIRLS 2021 international benchmark evidence for Austrian fourth-grade students is reported from two perspectives of the quality of primary education: the international median and trends from PIRLS 2016 as evidence of progress or decline.

3.2 Early childhood (Target 4.2)

In recent years, the importance of attending pre-primary school has gained momentum. The PIRLS Learning to Read Survey (or Parent Questionnaire) assesses the availability of these kinds of facilities but also interrogates the kinds of early reading activities at home that parents offer their children (IEA, 2016). The target focuses on both access to and completion of schooling, with indicators that emphasize literacy in reading and mathematics (UN, 2015). The study, however, only focuses on reading literacy through the analysis of PIRLS.

3.3 Skills for work (Target 4.4)

In addition to PIRLS, other large-scale international assessment results consistently show a gender difference for boys and girls in literacy, numeracy, and science, as evidenced by the Trends in International Mathematics and Science Study (TIMSS). PIRLS data highlight the need for systemic intervention to ensure equitable skills for work and also point to differences in exposure to digital resources and technology, ensuring relevant work skills for different labor market contexts (IEA, 2016). One argument can be that digital skills development fosters critical thinking and literacy in digital contexts, which are increasingly part of literacy (e.g., digital literacy) (Kong, 2014).

3.4 Equity (Target 4.5)

PIRLS allows for greater disaggregation of data to provide essential evidence for targeted intervention, monitoring, and planning for crucial sub-groups of the population who may be at a continued educational disadvantage (IEA, 2016). Equity targets for which data is collected by PIRLS include gender, language (for the current analyses, interpreted as the language most frequently spoken at home), and home socioeconomic status. Socioeconomic equity (e.g., access to resources, language spoken at home) affects literacy (Hemmerechts et al., 2017).

3.5 Learning environment (Target 4.a)

PIRLS provides information on the quality of the learning environment in terms of bullying, school safety, and factors that impede teaching practice. Student questionnaires gauge attitudes, opinions, and instances of bullying and the severity thereof, as bullying may not only be a hindrance to academic performance and well-being in the early years but may well continue into secondary school and phases beyond the initial grades when firm foundations in a climate of safety and orderliness are of great importance. Teacher questionnaires gather information about school safety issues and associated factors that severely affect teacher ability to deliver the curriculum effectively (IEA, 2016). A safe and supportive learning environment can impact literacy outcomes by reducing distractions like bullying and creating conditions conducive to learning (Darling-Hammond & Cook-Harvey, 2018; De Nobile et al., 2017).

3.6 Teachers (Target 4.c)

The supply of a qualified teacher workforce is crucial for every education system. While a universally accepted definition of a qualified teacher cannot be applied across contexts and participating countries, PIRLS provides information on the highest levels of qualification obtained for each country (IEA, 2016). While formal qualifications provide some indication of who is needed in front of the classroom to ensure success, other indicators of pedagogical training (such as classroom language, reading pedagogy, reading theory, and assessment methods) refine the outcomes that are obtained over and above formal qualifications as the sole indicator of quality. Teacher qualifications might (or might not) influence literacy outcomes (Guo et al., 2012).

The PIRLS 2021 Assessment Framework is crucial in understanding literacy in relation to SDG 4 and exploring the connections between literacy and other educational goals. It serves as a reference point for investigating these intersections, leading to the research questions discussed next.

4 Research questions

Since Targets 4.2, 4.4, 4.5, 4.a, and 4.c are directly linked to the items in the PIRLS 2021 questionnaire and are plausible predictors of literacy (Target 4.1), the current study aims to investigate the extent of PIRLS 2021 evidence in operationalizing SDG 4 targets and indicators by asking the following questions:

1. What do overall benchmark achievements indicate about the state of primary education in Austria as measured by PIRLS 2021?
2. To what extent do early home literacy activities, as part of Target 4.2, shape reading literacy outcomes?
3. How does exposure to digital resources and technology, as part of skills for work in Target 4.4, affect reading achievement?
4. What role does equity (Target 4.5) play when home socioeconomic factors, language spoken at home, and gender are included as predictors of fourth-grade Austrian students' reading literacy achievement?
5. How are learning environments (in terms of bullying, safe and orderly schools, and factors that limit teacher practice) (Target 4.a) and teachers' highest levels of qualification (Target 4.c) related to reading literacy outcomes?
6. How are results from ILSAs meaningful for SDG 4 targets and indicators in expanding the understanding of these in interconnected ways?

The research questions enable a thorough analysis by evaluating various aspects such as direct measures of academic achievement, contextual factors that impact learning, and the alignment of the findings with the global educational goals set by SDG 4. Further discussion on how this analysis will be conducted is explained below.

5 Methods

5.1 Design

This study takes the form of a secondary analysis of PIRLS 2021 using Austrian data. As a developed country in central Europe, Austrian participation in PIRLS dates back to its first participation in the study in the 2001 cycle. PIRLS is administered to children in their fourth year of schooling. PIRLS 2021 placed Austrian fourth-grade student achievement at 530 (SE = 2.2), a score substantially above the PIRLS scale center point of 500. While these overall results for Austria are encouraging, there was a drop of 11 score points between PIRLS 2016 and PIRLS 2021. This decline may be due to the effects of COVID-19, since PIRLS 2021 was administered amid the school disruption and closure of the pandemic.

PIRLS 2021 consists of fourth-grade achievement data and contextual background data collected from school principals, fourth-grade teachers, and the parents (or primary caregivers) of fourth-grade students. When using teacher and parent data, results are reported regarding the teachers or parents of fourth-grade students since the results are representative at the student level, not the teacher or parent level.

5.2 Sample

A total of 160 Austrian schools participated in PIRLS 2021, and 4,806 fourth-grade students were assessed. From the sampled classes, 305 teachers completed the Teacher Questionnaire and 4,806 parents of fourth-grade students completed the Parent Questionnaire (referred to as the Learning to Read Survey).

5.3 Data collection

Target 4.1 aims to promote fair and just education outcomes by measuring the percentage of children and adolescents who achieve a particular level of proficiency in fundamental subjects at key educational milestones. Specifically, this target assesses student proficiency in reading, writing, and math (a) during the early grades, (b) after completing primary education, and (c) at the end of lower secondary education, with a detailed gender breakdown. This approach concentrates on significant stages in educational development, evaluating foundational learning that can pave the way for future academic success (Indicator 4.1.1). PIRLS 2021 tested fourth-grade student achievement using plausible values on four proficiency scales, namely the Low International Benchmark, the Intermediate International Benchmark, the High International Benchmark, and the Advanced International Benchmark. To provide evidence for research question 1, the PIRLS 2021 overall achievement is used to indicate reading achievement for fourth graders on each of the international benchmarks and also by gender. Since PIRLS does not provide data regarding completion rates, Indicator 4.1.2 was excluded from the analysis in the current study.

To address research questions 2, 3, 4, and 5, reading achievement data in the form of overall plausible values from fourth-grade students were used in conjunction with contextual background data from the Teacher and Parent Questionnaires. These questionnaires were administered to provide a comprehensive understanding of the factors that shape reading literacy outcomes in relation to several SDG 4 targets. Composite scales by means of sum scores were devised for each of the SDG 4 targets. The direction of these scales was all computed to mean that the higher the value, the more of a particular activity or characteristic was present.

Table 1 illustrates the composition of the SDG target scales using PIRLS 2021 contextual background variables to answer research question 2, which asked the extent to which early home literacy activities shape reading literacy outcomes.

Table 1
Target 4.2 – Early childhood and PIRLS 2021 variable composition

Target	Question wording in the PIRLS Early Learning Survey (Parent) Questionnaire	Variable name
Target 4.2: Early childhood	Did your child attend pre-primary school? Yes/No	ASBH05AB
	Before your child began primary/elementary school, how often did you or someone else in your home do the following activities with him or her? Often, Sometimes, Never or almost never a) Read books b) Tell stories c) Sing songs d) Play with alphabet toys (e.g., blocks with letters of the alphabet) e) Talk about things you had done f) Talk about what you had read g) Play word games h) Write letters or words i) Read aloud signs and labels	ASBH01A-I (ASBHELA: Early Literacy Activities scale)

Research question 3 asked about how exposure to digital resources and technology, as part of skills for work in Target 4.4, affects reading achievement. Table 2 indicates the variables used for measuring Target 4.4.

Table 2
Target 4.4 – Skills for work and PIRLS 2021 variable composition

Target 4.4: Skills for work	How much do you agree with these statements about using computers, tablets, or smartphones? a) I am good at using a computer or tablet b) I am good at typing c) It is easy for me to find information on the internet d) I know how to create written stories or reports e) I know how to create presentations f) I can recognize a website that is useful to me g) I can tell if a website is trustworthy h) I know how to make and share a video	ASBG09A-H (ASBGSEC: Digital self-efficacy scale)
	Question wording in the PIRLS Teacher and Student Questionnaire	Variable name
Target 4.4: Differences in exposure to digital resources and technology	What access do the students have to digital devices? Yes/No a) The school provides each student with a digital device b) The class has digital devices that students can share c) The school has digital devices that the class can use sometimes d) Students bring their own digital devices	ATBR12B

Target 4.5 deals with issues of gender, home socioeconomic status, and language. Table 3 indicates how scales for this target were devised in relation to research question 4.

Table 3
Target 4.5 – Equity and PIRLS 2021 variable composition

	Question wording in the PIRLS Student and Parent Questionnaire	Variable name
Target 4.5: Equity	Which best describes you? Girl Boy	ITSEX
	About how many books are there in your home? (Do not count e-books, magazines, newspapers, or children's books.) 0–10 11–25 26–100 101–200 More than 200 About how many children's books are there in your home? (Do not count children's e-books, magazines, or schoolbooks.) 0–10 11–25 26–50 51–100 More than 100 Do you have any of these things in your home? a) Access to the internet b) A computer, tablet, or e-reader c) A smartphone Highest level of education of either parent: Finished some primary or lower secondary or did not go to school 2) Finished lower secondary 3) Finished upper secondary 4) Finished post-secondary education 5) Finished university or higher Highest level of occupation of either parent: Has never worked outside the home for pay, general laborer, or semi-professional (skilled agricultural or fishery worker, craft or trade worker, plant or machine operator), 2) Clerical (clerk or service or sales worker), 3) Small business owner, 4) Professional (corporate manager or senior official, professional, or technician or associate professional)	ASBH12 (ASBHSES: Home Resources for Learning scale) ASBH13 ASDHEDUP ASDHOCCP
	How often do you speak German at home? Always, Almost always, Sometimes, Never	ASBH04

To answer research question 5 of the current study, students and teachers were asked about the extent to which the learning environment (in terms of bullying, safe and orderly schools, and factors that limit teacher practice) and teachers' highest levels of qualification are related to reading literacy outcomes. Target 4.a is operationalized in Table 4.

Table 4
Target 4.a – Learning environment and PIRLS 2021 variable composition

	Question wording in the PIRLS Student and Teacher Questionnaire	Variable name
Target 4.a: Learning environment	During this year, how often have other students from your school done any of the following things to you, including through texting or the internet? a) Made fun of me or called me names. b) Left me out of their games or activities. c) Spread lies about me. d) Stole something from me. e) Damaged something of mine on purpose. f) Hit or hurt me (e.g., shoving, hitting, kicking). g) Made me do things I didn't want to do. h) Sent me nasty or hurtful messages online. i) Shared nasty or hurtful information about me online. j) Threatened me.	ASBG11A-J (ASBGSB: Student Bullying scale)
	Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements: a) This school is located in a safe neighborhood. b) I feel safe at this school. c) This school's security policies and practices are sufficient. d) The students behave in an orderly manner. e) The students are respectful of the teachers. f) The students respect school property. g) This school has clear rules about student conduct. h) This school's rules are enforced in a fair and consistent manner. i) The students are respectful of each other,	ATBG11A-I (ATBGSOS: Safe and Orderly School scale)
	In your view, to what extent do the following limit how you teach this class? a) Students lacking prerequisite knowledge or skills b) Students suffering from lack of basic nutrition c) Students suffering from not enough sleep d) Students absent from class e) Disruptive students f) Uninterested students g) Students with mental, emotional, or psychological impairment h) Students needing extra support in reading	ATBR03A-H (ATBGSLI: Classroom Teaching Limited by Students Not Ready for Instruction scale)

Lastly, Target 4.c, which deals with teacher qualifications, was operationalized by the PIRLS 2021 Teacher Questionnaire data that asked teachers about their highest level of formal qualification. Table 5 indicates the response options for teachers' highest levels of qualification.

Table 5
Target 4.c – Teachers and PIRLS 2021 variable composition

	Question wording in the PIRLS Teacher Questionnaire	Variable name
Target 4c: Teachers	What is the highest level of formal education you have completed? Did not complete/Upper secondary education – ISCED level 3 Upper secondary education – ISCED level 3 Post-secondary/non-tertiary education – ISCED Level 4 Short-cycle tertiary education – ISCED Level 5 Bachelor's degree or equivalent level – ISCED Level 6 Master's degree or equivalent level – ISCED level 7 Doctoral degree or equivalent level – ISCED level 8	ATBG04

After collecting data from PIRLS 2021 and SDG 4 targets (UN, 2015), we utilized quantitative techniques to extract meaningful insights and patterns from the dataset.

5.4 Data analysis

All data were analyzed using the International Database (IDB) Analyzer, which is software specifically developed to analyze large-scale international data with SPPS as the operating platform. To answer research question 1, the overall benchmark results for Austrian fourth-grade children are presented as reported in the PIRLS 2021 International Report (see Mullis et al., 2023). To address research questions 2–5, linear regression was performed on data from the parent, teacher, and student background questionnaires to determine the extent of the possible effect of a number of scales on overall fourth-grade Austrian reading achievement.

6 Results

PIRLS 2021 provides reading achievement results for each of the four international benchmarks. These benchmarks represent four levels of reading competence. Figure 3 illustrates the differences between the international median for PIRLS 2021 International Benchmarks and the Austrian fourth-grade student benchmark achievement:

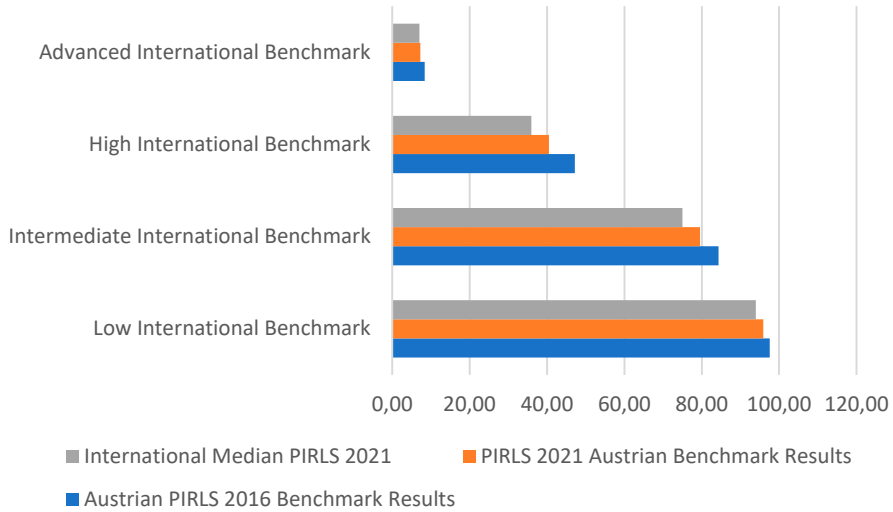


Figure 3
International Benchmarks and the Austrian fourth-grade student benchmark achievement in PIRLS 2016 and PIRLS 2021.

Austrian fourth-grade achievement (represented in Figure 3by the orange bars for PIRLS 2016 and blue bars for PIRLS 2021) at the Advanced International Benchmark resembles percentages of students internationally who were able to reach this benchmark (as indicated by the gray bar in Figure 3). Higher percentages of Austrian students were able to reach the high, intermediate, and low benchmarks, respectively, than their international counterparts. While these signals are good indicators of the Austrian education system, PIRLS 2021 results have shown a slight decrease from PIRLS 2016 benchmark results (Mullis et al., 2017). The effects of COVID-19 could have played a significant role in these decreases, as can be seen from overall trend results from other European countries: France and Turkey showed increases in reading achievement between PIRLS 2016 and PIRLS 2021; decreases were found for Belgium (from 497 to 494 score points), Denmark (from 547 to 539 score points), Germany (from 537 to 524 score points), and Italy (from 548 to 537 score points) among others (Mullis et al., 2023). Nonetheless, Austrian fourth-grade benchmark achievement was maintained at and above the international median in PIRLS 2021 despite overall trend score decreases.

Table 6
Regression results

	b	*SE* B	*t*
(CONSTANT)	188.33	19.5	
Target 4.2 Early childhood development	1.7	0.7	4.4
Target 4.4 Skills for work	1.5	0.7	2.1
Target 4.5 Equity	14.0	0.5	28.1
Target 4.a Learning environment	2.9	0.4	6.8
Target 4.c Highest level of qualification	0.2	4.6	0.5

Table 6 provides the results of the regression analysis for SDG targets that are specifically addressed by the PIRLS 2021 data. The current model explains 32% of the variance in the data, with reliable scales constructed for early childhood development (0.71), and skills for work as measured by the digital self-efficacy scale (0.75). Items that measure exposure to digital devices (see Table 2) were removed from the analysis due to low reliability coefficients. Cronbach's alpha values for the Learning Environment – bullying (0.87), school safety (0.85), and classroom environment (0.76) – were all within acceptable ranges. Home socioeconomic status, as part of the equity target, is a formative construct; calculation of Cronbach's alpha is conceptually meaningless (Stadler et al., 2021).

Of statistical significance in relation to overall reading achievement were those targets related to early childhood development, skills for work equity. and the learning environment, with t-values larger than 1.96 at the 0.05 level of significance. Good early childhood practices seem to make a substantial difference: 97% (*SE* = 0.3) of children who attended pre-primary schools (of whom approximately 80% attended for three years or more) and had parents who engaged in early literacy activities clearly benefited. More than two-thirds of parents of fourth-grade students reported that they often engage their children in reading books, talking about what they had done, counting things, playing games with shapes, and using building blocks. There seems to be room for improvement: 61.5% (*SE* = 0.9) of parents of fourth-grade students reported that their children could not complete early literacy tasks very well. These tasks included recognizing letters of the alphabet, reading some words, sentences, and a story, as well as writing letters of the alphabet and some words. Where literacy activities could be translated into tasks, the expectation may be that the effect on reading literacy achievement might be bigger.

Children's digital self-efficacy as a proxy for skills at work shows a significant effect on reading achievement for the majority of fourth-grade students who indicated high (39.4%, *SE* = 1.0) and moderate (47.5%, *SE* = 1.0) digital self-efficacy.

Equity has a substantial effect of 14.0 score points more in those environments of higher home socioeconomic status, which is true for more than three-quarters of the parents of fourth-grade Austrian students who participated in PIRLS 2021. Gender as part of the equity target provides statistically significant evidence that boys still trail behind girls by as much as 9.8 score points in overall reading achievement. For 74.4% (*SE* = 0.9) of Austrian students, German is most often spoken at home.

A positive learning environment is statistically significantly associated with reading literacy achievement as indicated by 61.2% (*SE* = 1.0) of fourth-grade students who reported that they had never been bullied. A majority of teachers experience the learning environment as very safe (51.63%, *SE* = 3.2) or somewhat safe (47.81%, *SE* = 3.2); their assessment of their teaching being limited by students who are not ready was as low as 2.5% (*SE* = 1.0). While questions to teachers about access to digital devices did not form a psychometrically sound scale, overall descriptive statistics show that the difference between teachers of fourth-grade students who reportedly use a variety of digital devices in a range of ways only once a week (39.2%, *SE* = 3.5) is negligible and one score point higher when compared to those teachers of fourth-grade students who reportedly use these devices every day or almost every day (60.8%, *SE* = 3.5).

Lastly, while not statistically significant, it is interesting to note that teachers' formal qualifications do not show a linear relationship with overall reading achievement. Achievement for students whose teachers have Master's degrees can be expected to be higher by only 0.2 (SE04.9) points.

7 Discussion

The findings of PIRLS Austria 2021, as used in the current research, hold significant meaning for quality education as it can help support the political, economic, technical, and sociocultural aspects (Addey & Sellar, 2019). Developing reading literacy is crucial for an individual's involvement in the lifelong learning process, which supports the social and economic aspects of their life and plays a significant role in developing communities and nations (Priyadarshini, 2019). This study shows how SDG 4 targets are interconnected with each other and PIRLS data, as discussed in the following paragraphs.

Reading achievement results by the international benchmarks for PIRLS 2016 and PIRLS 2021 in Austria provided evidence that minimum reading proficiency levels are on track. While PIRLS benchmark results have shown a slight decrease between PIRLS 2016 and PIRLS 2021, Austrian fourth-grade benchmark achievement was maintained at and above the international median in PIRLS 2021. This study revealed a gender gap

in reading achievement, with boys' reading achievement falling substantially behind that of girls. This evidence highlights the need for equity-focused educational interventions to address such disparities and the essence of Indicator 4.5.1 regarding parity indices focusing on how boys could be empowered to improve their reading skills. Target 4.5 aims to eliminate gender disparities and ensure that everyone, including vulnerable populations, has equal access to education (UN, 2015).

The regression results (Table 6) highlight the importance of digital self-efficacy, early childhood education, a safe learning environment, and addressing equity concerns through a more comprehensive assessment of home socioeconomic factors. These results indicate the essence of aiming to achieve SDG 4 Targets 4.2 (equity in early childhood education), 4.5 (eliminating gender disparities), and 4a (having a safe learning environment) (UN, 2015). These targets are also related to the results in overall reading achievement, which also indicate a direct link to Targets 4.2, 4.5 and 4.a.

Early childhood findings highlight the fact that children who attended pre-primary school for three years or more and engaged in early literacy activities with their parents significantly show the importance of quality early childhood education. However, the attainment of certain early literacy tasks (e.g., recognizing letters of the alphabet), could enhance the role of the early education environment in addition to the role played by activities of interaction and play. This finding aligns with Target 4.2, which seeks to ensure access to quality early childhood development, care, and pre-primary education to prepare children for primary school in fair and equitable ways. When children feel respected and valued in their environment, they are more likely to thrive and reach their full potential (Banks, 2023).

Home socioeconomic status and frequently speaking the test language at home significantly affect reading scores. Research by Hemmerechts et al. (2017) and Beeharry (2021) highlighted the significant gaps in education and literacy, particularly in low-income countries where parental literacy involvement varies across socioeconomic status levels. The study by Van Staden, Bosker, and Bergbauer (2016) on prePIRLS 2011 South African data revealed that African children are significantly disadvantaged if they lack a strong foundation in their native language and receive education in a non-native language during their first three years of school. Exposure to a language that is at least similar to the home language can boost a child's reading performance (Van Staden et al., 2016), thereby highlighting the significance of speaking and being familiar with the language of the test at home to achieve better results.

The current research includes those aspects of the social learning environment at school in terms of safety and bullying. The fourth target's focus on safe, non-violent, inclusive, and effective learning environments

also includes the role of the home environment in education, particularly for early childhood learning (UN, 2015). This statement implies that learning environments are not limited to school premises. It is indeed essential to focus on learning environments in the school, but a broader goal has to be achieved in ensuring a safe learning environment overall, linking this finding to SDG 11: making cities and human settlements inclusive, safe, resilient, and sustainable.

The findings regarding digital self-efficacy speak to practically implementable interventions from as early as possible to ensure that children develop and hone their digital skills in anticipation of the world of work. Target 4.4 of the education agenda is closely related to skills for work and the use of digital resources. Positive digital self-efficacy from as early as possible points to the importance of introducing early interventions when these skills are not developing along a trajectory that would ensure adequate digital readiness for the world of work. Descriptive results show that there were no significant differences in reading achievement based on the frequency of digital device use by teachers. This finding might suggest that digital resources are not important for education. However, this finding highlights the need for a more nuanced understanding of how digital tools are integrated into learning.

These findings suggest that simply increasing access to digital devices, as aimed at in Target 4.4, may not be sufficient. Effective technology integration in education requires quality digital content, pedagogical strategies, and teacher training to improve educational outcomes (Eden et al., 2024). Kong (2014) showed how activities like accessing digital resources, processing information, and engaging in peer discussions improved student ability to critically analyze and synthesize information. SDG 9, which aims to increase access to technology, including (ICT), and to strive for universal and affordable internet access in the least developed countries by 2020 is only realistic when early literacy and skills regarding technology at the school level can be leveraged as tools for effective learning (Neumann, 2018; Paul et al., 2023; Sarker et al., 2019).

In pursuing quality education and lifelong learning for all, the impact of teacher qualifications on student achievement has been a subject of much debate and discussion (Antony & Elangkumaran, 2020). Although not statistically significant, evidence from this study suggests that having higher formal qualifications is not associated with higher reading achievement among students. While teacher formal qualifications are a basic indicator of quality that could be expected, other factors such as teacher attitude, motivation, self-efficacy teaching methods, interactions between teachers and students, and the learning environment may significantly influence student outcomes (Darling-Hammond & Cook-Harvey, 2018; De Nobile et al., 2017; Guo et al., 2012; Osman & Warner, 2020). This observation underlines the need for

a more comprehensive approach to teacher development, which would involve formal education and professional growth opportunities that improve teacher ability to engage and motivate students effectively.

The findings from the current study confirm that PIRLS provides valuable measures for monitoring and operationalizing the SDG 4 targets. Moreover, the current study's contribution and value extend beyond confirming correlations and statistically significant associations: they also indicate that SDG 4 targets are interconnected with one another and that ILSA data (as illustrated here by using PIRLS 2021 Austrian data) provide essential conceptual and empirical basis from which these conclusions are drawn. SDG 4 and its targets do not function in isolation from other SDGs. Therefore, the approach to reaching SDG targets should bear in mind that SDGs function with one another in an interconnected way. For example, the achievement of SDG 4 is closely linked with the attainment of SDG 1 (no poverty), SDG 3 (good health and well-being), SDG 5 (gender equality), SDG 8 (decent work and economic growth), and SDG 10 (reduced inequalities) (Nikolova & Suleimenova, 2019).

8 Limitations and recommendations

Although this research provides valuable insights, it has certain limitations that need to be acknowledged. First, ILSAs, like PIRLS, have a narrow focus on specific areas of literacy and numeracy that may overlook other crucial aspects of quality education, including critical thinking, creativity, and socio-emotional skills. These skills are also essential components of SDG 4, and their exclusion may lead to an incomplete understanding of educational quality.

Second, there are certain limitations when measuring progress towards SDG 4, including underlying gender-based disparities in educational outcomes. It is recommended that more research be conducted on these disparities. Additionally, PIRLS data does not capture some SDG 4 indicators, such as completion rates. Therefore, it is suggested to incorporate data from alternative sources that track completion rates and other relevant indicators. These sources could include national educational statistics, reports from UNESCO, or other international assessments that include broader educational metrics. A multi-source approach can better evaluate all SDG 4 targets and enhance the overall understanding of educational progress.

Lastly, the data analysis approach has limitations. As with any secondary analysis, the available data are utilized without the advantage of adding other data that may be relevant to the analyses. It must also be noted that the use of linear regression methods only establishes relationships between predictor and outcome variables, not causality.

9 Conclusion

This research highlights the vital role of SDG 4 in achieving the remaining SDGs by demonstrating its interconnectedness with other aspects of the SDG agenda. Several sources, including Madalinska-Michalak (2023), OECD (2017), Priyadarshini (2019), UNESCO (2021), and the UN (2018) have supported this notion. This research emphasizes that reading comprehension and early childhood development are essential for literacy as an outcome of quality education. Additionally, empirical data sources play a significant role in establishing systemic relations between role players. The study further emphasizes that ILSAs, as the body of empirical evidence, not only support the targets set by the SDGs but also make these targets measurable and concrete. This research shows that SDG 4 and ILSAs like PIRLS play a crucial role in providing empirical evidence for quality education and are important for achieving other SDGs. The study also highlights the need for collective action to ensure everyone has access to safe learning opportunities, quality education, and improvement of literacy from an early age.

References

Addey, C., & Sellar, S. (2019). Is it worth it? Rationales for (non) participation in international large-scale learning assessments. *Education Research and Foresight: Working Papers, 24*. https://unesdoc.unesco.org/ark:/48223/pf0000368421

Antony, S., & Elangkumaran, P. (2020). An impact on teacher qualifications on student achievement in science: A study on the G.C.E. (O/L) in Trincomalee District. *International Journal of Engineering Science and Computing, 10*(2), 24690-24695.

ASviS. (2019). *The European Union and the sustainable development goals*. ASviS. https://asvis.it/public/asvis2/files/Pubblicazioni/Compositi_Europei_ENG_HR.pdf

Banks, B. (2023, 20th November). *Why equity is key to quality in early childhood education*. Childcare Aware of Kansas. https://ks.childcareaware.org/equity-in-childhooded/

Beeharry, G. (2021). The pathway to progress on SDG 4 requires the global education architecture to focus on foundational learning and to hold ourselves accountable for achieving it. *International Journal of Educational Development, 82*, 1–7. https://doi.org/10.1016/j.ijedudev.2021.102375

Bruns, B., Macdonald, I. H., & Schneider, B. R. (2019). The politics of quality reforms and the challenges for SDGs in education. *World Development, 118*, 27–38.

Dahl, A. L. (2012). Achievements and gaps in indicators for sustainability. *Ecological Indicators, 17*, 14–19. https://doi.org/10.1016/j.ecolind.2011.04.032

Darling-Hammond, L., & Cook-Harvey, C. M. (2018). *Educating the Whole Child: Improving School Climate to Support Student Success*. Learning Policy Institute.

De Nobile, J., Lyons, G., & Arthur-Kelly, M. (2017). *Positive learning environments: Creating and maintaining productive classrooms*. Cengage AU.

Eden, C. A., Chisom, O. N., & Adeniyi, I. S. (2024). Harnessing technology integration in education: Strategies for enhancing learning outcomes and equity. *World Journal of Advanced Engineering Technology and Sciences, 11*(2), 1–8.
https://doi.org/10.30574/wjaets.2024.11.2.0071

Eurostat. (2022). *Sustainable Development in the European Union. Monitoring Report on Progress towards the SDGs in an EU Context – 2022 edition*. Eurostat.
https://ec.europa.eu/eurostat/web/products-flagship-publications/-/ks-09-22-019

Fehlner, W. (2019). Educating for sustainability: The crucial role of the tertiary sector. *Journal of Sustainable Development*, 12(2), 18–28.

Flores, M. A. (2023). Unpacking quality in teacher education. In J. Madalinska-Michalak (Ed.), *Quality in teaching and teacher education: International perspectives from a changing world* (pp. 32–49). Brill.
https://doi.org/10.1163/9789004536609_003

Grobler, S. & Dittrich, A.-K. (2024). Envisioning quality education for sustainability transformation in teacher education: perspectives from an international dialogue on Sustainable Development Goal 4. *International Journal of Comparative Education and Development*, *26*(3).
https://doi.org/10.1108/IJCED-06-2023-0048

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.

Havea, P. H., & Mohanty, M. (2020). Professional development and sustainable development goals. In W. Leal Filho, A. M. Azul, L. Brandli, P. G. Özuyar, & T. Wall, *Quality Education* (pp. 654 – 665). Springer.

Hemmerechts, K., Agirdag, O., & Kavadias, D. (2017). The relationship between parental literacy involvement, socio-economic status and reading literacy. *Educational Review*, *69*(1), 85–101.
https://doi.org/10.1080/00131911.2016.1164667

IEA (2016). *Measuring SDG 4: How PIRLS can help*. IEA.
https://www.iea.nl/sites/default/files/2019-04/SDG4_targets_PIRLS_2016_Infographics_EN.pdf (iea.nl)

IEA. (2021). *PIRLS 2021 encyclopedia: Education policy and curriculum in reading*. TIMSS & PIRLS International Study Center.
https://pirls2021.org/download

Kong, S. C. (2014). Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy. *Computers & Education*, *78*, 160–173.

Kumar, C. (2020). Scope of education for sustainable development. In W. Leal Filho, A. M. Azul, L. Brandli, P. G. Özuyar, & T. Wall, *Quality Education* (pp. 741–752). Springer.

Leino, K., Nissinen, K., & Sirén, M. (2022). Associations between teacher quality, instructional quality and student reading outcomes in Nordic PIRLS 2016 data. *Large-Scale Assessments in Education, 10*, Article 25.
https://doi.org/10.1186/s40536-022-00146-4

Ledger, S., Thier, M., Bailey, L., & Pitts, C. (2019). OECD's approach to measuring global competency: Powerful voices shaping education. *Teachers College Record, 121*(8), 1−40.
https://doi.org/10.1177/016146811912100802

Madalinska-Michalak, J. (2023). Introduction. In J. Madalinska-Michalak (Ed.), *Quality in teaching and teacher education: International perspectives from a changing world* (pp. 1–8). Brill.
https://doi.org/10.1163/9789004536609_001

Moldan, B., & Dahl, A. L. (2007). Challenges to sustainability indicators. In T. Hák, B. Moldan, & A. L. Dahl (Eds.), *Sustainability indicators* (pp. 1–26). Island Press.

Mullis, I. V.S., Martin, M. O., Minnich, C. A., Drucker, K. T., & Ragan, M. A. (2012). *PIRLS 2011 encyclopedia: Education policy and curriculum in reading.* International Association for the Evaluation of Educational Achievement (IEA).

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). What makes a good reader: International findings from PIRLS 2016. *PIRLS 2016 International results in reading.* TIMSS & PIRLS International Study Center.
https://pirls2016.org/pirls/summary/index.html

Mullis, I. V. S, & Martin, M. O., (Eds.). (2019). *PIRLS 2021 assessment frameworks.* TIMSS & PIRLS International Study Center.
https://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf

Mullis, I. V. S., Von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021 international results in reading.* TIMSS & PIRLS International Study Center.
https://doi.org/10.6017/lse.tpisc.tr2103.kb5342

Neumann, M. M. (2018). Using tablets and apps to enhance emergent literacy skills in young children. *Early Childhood Research Quarterly, 42*, 239–246.
https://doi.org/10.1016/j.ecresq.2017.10.006

Nikolova, A., & Suleimenova, Z. (2019). *Introduction to the 2030 Agenda and the interconnectedness of the SDGs.* UNESCAP/UNDP Training Workshop, 26–28 February.
https://www.unescap.org/sites/default/files/Introduction%20to%20the%202030%20Agenda%20and%20the%20Interconnected%20SDGs_Eng.pdf

OECD. (2017). *Education at a glance 2017: OECD indicators.* OECD.
https://doi.org/10.1787/eag-2017-en

OECD. (2022). *The Short and Winding Road to 2030: Measuring Distance to the SDG Target.*
https://www.oecd.org/en/publications/the-short-and-winding-road-to-2030_af4b630d-en.html

Osman, D. J., & Warner, J. R. (2020). Measuring teacher motivation: The missing link between professional development and practice. *Teaching and Teacher Education, 92*, Article 103064.
https://doi.org/10.1016/j.tate.2020.103064

Paul, C. D., Hansen, S. G., Marelle, C., & Wright, M. (2023). Incorporating technology into instruction in early childhood classrooms: A systematic review. *Advances in Neurodevelopmental Disorders, 7*(3), 380–391.
https://doi.org/10.1007/s41252-023-00316-7

Priyadarshini, A. (2019). Evolving literacy perspectives: Towards lifelong learning and sustainable development. In W. Leal Filho, A. M. Azul, L. Brandli, P. G. Özuyar, & T. Wall, *Quality Education* (pp. 295–306). Springer.
https://doi.org/10.1007/978-3-319-69902-8_42-1

Robinson, K., & Aronica, L. (2016). *Creative schools: The grassroots revolution that's transforming education*. Penguin Books.

Sayed, Y., & Moriarty, K. (2020). SDG 4 and the 'Education Quality Turn': Prospects, possibilities, and problems. In A. Wulff (Ed.), *Grading Goal Four: Tensions, threats, and opportunities in the Sustainable Development Goal on quality education* (pp. 194–213). Brill.
https://doi.org/10.1163/9789004430365_009

Sarker, M. N. I., Wu, M., Cao, Q., Monirul Alam, G. M., & Li, D. (2019). Leveraging digital technology for better learning and education: A systematic literature review. *International Journal of Information and Education Technology, 9*(7), 453–461.
https://www.ijiet.org/vol9/1246-JR329.pdf

Schleicher, A. (2017). The future of PISA. *Tertium Comparationis*, *23*(1), 113–125.

SDSN & IEEP. Europe Sustainable Development Report 2021. (2021). *Transforming the European Union to Achieve the Sustainable Development Goals*. Pica Publishing.
https://www.pica-publishing.com.

Stadler, M., Sailer, M., & Fischer, F. (2021). Knowledge as a formative construct: A good alpha is not always better. *New Ideas in Psychology, 60*, 100832.

Statistik Austria. (2022). *Bildung in Zahlen. Tabellenband*. Statistik Austria.
https://www.statistik.at/fileadmin/pages/325/Bildung_in_Zahlen_20_21_Tabellenband.pdf

Tatto, M. T., & Pippin, J. (2017). The quest for quality and the rise of accountability systems in teacher education. In J. Husu, & D. J. Clandinin (Eds.), *The Sage handbook of research on teacher education* (pp. 68–89). Sage.
https://doi.org/10.4135/9781526402042.n4

Thangeda, A., Baratiseng, B., & Mompati, T. (2016). Education for Sustainability: Quality Education Is a Necessity in Modern Day. How Far Do the Educational Institutions Facilitate Quality Education? *Journal of Education and Practice*, *7*(2), 9–17.

Thaung, N. N. (2018, October 15-19). *Monitoring of SDG4: Global and regional level*. Regional Training Course on Education-Related SDG indicators.
https://www.unsiap.or.jp/face-to-face/pss/es/1810_kor.html

Trane, M., Marelli, L., Siragusa, A., Pollo, R., & Lombardi, P. (2023). Progress by research to achieve the sustainable development goals in the EU: a systematic literature review. *Sustainability*, *15*(9), 7055.
https://doi.org/10.3390/su15097055

UN. (2015, November 6). *Transforming our world: The 2030 agenda for sustainable development*. United Nations General Assembly.
https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

UN. (2018). *Quality education: Why it matters*. United Nations.
https://www.un.org/sustainabledevelopment/wp-content/uploads/2017/02/4_Why-It-Matters-2020.pdf

UNESCO. (2015). *Education for All 2000–2015: Achievements and Challenges*. UNESCO.

UNESCO. (2017). *Unpacking Sustainable Development Goal 4 Education 2030: Guide*. UNESCO. https://docs.campaignforeducation.org/post2015/SDG4.pdf

UNESCO. (2023, April 20). *UNESCO: Sustainable development begins with education*. UNESCO Press release. https://www.unesco.org/gem-report/en/articles/unesco-sustainable-development-begins-education

Unterhalter, E. (2019). *The many meanings of quality education: Politics of targets and indicators in SDG4*. *Global Policy*, *10*(1), 39–51. https://doi.org/10.1111/1758-5899.12591

Van Staden, S., Bosker, R., & Bergbauer, A. (2016). Differences in achievement between home language and language of learning in South Africa: Evidence from prePIRLS 2011. *South African Journal of Childhood Education, 6*(1). https://dx.doi.org/10.4102/sajce.v6i1.441

Van Staden, S., & Schreiner, C. (2023). Shaping national use of international large-scale assessment data: Translating descriptive patterns into meaningful contextual scales of teacher reading practice. In S. van Staden & C. Combrinck (Eds.), *Tracking changes in South African reading literacy achievement: A developing context perspective* (pp. 121–150). Brill. https://doi.org/10.1163/9789004687011_006

Waldow, F. (2017). Projecting images of the 'good' and the 'bad school': Top scorers in educational large-scale assessments as reference societies. *Compare: A Journal of Comparative and International Education*, *47*(5), 647–664. https://doi.org/10.1080/03057925.2016.1262245

World Bank. (2021). *Countries and economies*. World Bank. https://data.worldbank.org/country

STUDIA PAEDAGOGICA

vol. 29, n. 3, 2024
https://doi.org/10.5817/SP2024-3-4

STUDY

ANALYZING STUDENT PERFORMANCE IN CONNECTION WITH LEARNING DISORDERS AND EXTRACURRICULAR ACTIVITIES IN FOREIGN LANGUAGE CLASSES

Roland Hegedűs[a], Krisztina Sebestyén[b]

[a] Faculty of Pedagogy, Catholic University of Károly Eszterházy
[b] Faculty of Pedagogy, Gál Ferenc University

ABSTRACT

In our study, we investigate the academic performances of students with atypical developmental pathways (dyslexia, dysgraphia, and dyscalculia) as compared to students without these disorders. Specifically, we consider how their performance is influenced by participation in classes with varying curricula and by their family background. To do this, we analyze data from the Hungarian National Competency Measurement 2019, tenth grade (83,751 students), of whom 1,515 students had one or more learning disorders (dyslexia, dysgraphia, dyscalculia). The students were enrolled in the standard curriculum, in intensive foreign language learning experiences, or in an increased number of lessons. We conducted a secondary analysis of the SPSS by creating two- and three-dimensional ANOVA tables and cross-tabulation analysis. Family background is examined using a central complex variable created from several variables by the Hungarian Education Office. Our results show that learning disabled students enrolled in a foreign language class have higher competence performance scores in both reading comprehension and mathematics than learning disabled students enrolled in a standard curriculum class, but lower than those enrolled in a class with an increased number of lessons; the same pattern is seen in participation in extracurricular activities. It seems that higher-achieving students enroll in advanced level foreign language training, but participants in advanced level training in other subjects are able to achieve even higher results, which may be due to family background. Different learning disorders affect student performance, but overall, the order set up based on student performance in classes with different curricula is the same for students with typical and atypical development.

KEYWORDS
specific learning disorder; student performance; family background; extracurricular activities; secondary school; National Competency Measurement; Hungary

CORRESPONDING AUTHOR
Roland Hegedűs, Faculty of Pedagogy, Catholic University of Károly Eszterházy, Klapka György út 12, H-3300 Eger, Hungary
e-mail: hegedusroland1989@gmail.com

Introduction

There is a wide body of literature on student performance, but it has tended to focus on the performance of students with typical development; little research has been conducted on students with atypical development (LeRoy et al., 2019). Atypical development includes several subcategories. In the present study, we focus on students with dyslexia, dysgraphia, dyscalculia, or a combination of these specific learning disorders. In Hungary, these students constitute almost half of the special educational needs category, and they participate in integrated, inclusive education alongside their peers with typical development (Hegedűs, 2023). We chose these groups for the study because they have a problem with one aspect of their learning, and we wanted to see how this affects their academic performance. Students with and without learning disorders are in the same class and have to meet the same requirements, so the aim of our study is to compare their performances to see how well they are doing at the required level. To assess their performance, we use the results of the National Competency Measurement (NCM), which includes their mathematics and reading comprehension results.

Learning a foreign language or a second language requires a different way of thinking than was used for learning the first language, and this can present a particular challenge for students with atypical development. Therefore, in our study, we are mainly interested in the outcomes of typical and atypical learners in classes with intensive foreign language learning. In order to interpret their results, we will also examine the results of students in standard curriculum classes and those in classes with an increased number of lessons in other subjects.

We look at which types of classes – with a standard curriculum or with an increased number of lessons – students do best in, and whether extra lessons and family background really have a positive impact on their performance. Examining family background is important both because it is an even stronger

determinant of student performance in Hungary than in other countries and because it also affects whether students can participate in school-only or extracurricular activities, the latter of which may be more likely to be paid activities (Hegedűs, 2020; Róbert, 2004). Research has suggested that participation in extracurricular activities could be important for people with learning disorders, for example, because the experiences and successes they have in such activities can have a positive impact on their academic performance, which is otherwise lower due to their learning disorder (Hegedűs, 2024).

There is a limited body of literature on the achievement of students with learning disorders. The description of the Hungarian situation and the factors influencing the learning outcomes of Hungarian students may be of interest internationally because what affects the performance of students with learning disorders in Hungary may also be present in other countries. Factors that have a positive impact in Hungary may also have an impact in other countries, thereby increasing the achievement of students in those countries. Furthermore, the study is significant because studies on whole populations are rarely done on such a large sample, so the results here can be considered well supported.

1 A brief introduction to specific learning disorders

A specific learning disorder can be considered a sub-disorder because students with these disorders show significant deficits in basic cultural techniques (writing, reading, counting), but no problems in other areas; they may even be highly gifted in some areas (Beckmann & Minnaert, 2018; Englbrecht & Weigert, 1996; Kirk & Bateman, 1962). The development of specific learning disorders is most often due to neurological or neuropsychological causes (e.g., neurodevelopmental dysfunction), hereditary factors, negative environmental influences, and external influences, but the specific cause is still unknown (Beckmann & Minnaert, 2018). The literature is not consistent on the relationship between intelligence level and specific learning disorders, with some researchers suggesting that these disorders develop at low IQ levels (75-85 IQ) (Englbrecht & Weigert, 1996); more recent literature has suggested that these disorders are more likely to develop above 90 IQ (Mesterházi & Szekeres, 2019; Şahin et al., 2020).

One specific learning disorder is dyslexia (reading disorder); dyslexia may develop even during growth in the womb, as a result of adverse and/or neurobiological abnormalities, but it does not affect the full spectrum of intelligence (Gyarmathy, 1998, 2018). Acquired dyslexia usually occurs later in life due to various brain trauma, head injury, stroke, infection, or cancer (Chase & Tallal, 1992; Gyarmathy, 2018). Dyslexia can have a number of symptoms, such as impaired reading sub-tasks (word recognition, oral reading

skills) or reading comprehension (BNO, 2004). The neurological abnormality in dyslexia is unquestionable, and it is likely to be a problem in the connectivity system in the brain network (Csépe, 2008; Žabkar et al., 2023). Dyslexia is often associated with spelling disorders and, as secondary symptoms, behavioral and emotional disorders (Cristofani et al., 2023).

Dysgraphia is also a specific learning disorder, a disorder of writing execution, namely a problem with the technical ability to write; dysgraphia can present symptoms that vary depending on gender and age (Biotteau et al., 2019). We also distinguish dysorthography, which is a disorder in the application of spelling rules (Dékány & Mohai, 2012). According to Gerebenné Várbíró (1995), dysgraphia is of neurological origin and can be caused by an underdevelopment of movement, spatial and temporal orientation, vision, hearing, speech hearing, and the visual system. Symptoms of dysgraphia may include slower writing speed, difficulty in recalling the shape of letters, poor verbal memory, errors in marking duration, using capital letters, spaces and punctuation, fine motor and graphomotor errors, incorrect spelling, and problems with line spacing (Chordia et al., 2020; Chung et al., 2020; Meixner, 2015). Dysgraphia is often associated with other specific learning disorders such as dyslexia (Chung et al., 2020).

Specific learning disorders include dyscalculia, a disorder of arithmetic skills that may appear during development or later as a result of brain injury (Márkus, 2007). The literature is not consistent on what underlies dyscalculia: damage to neuropsychological structures and functions (Farkasné Gönczi, 2011; Jiménez-Fernández, 2016), which does not affect general intelligence, or damage to mathematical solution systems, such as spatial-visual impairment (Rourke, 1993). Dyscalculia is a disorder of basic skills such as addition, subtraction, multiplication, and division; it has less effect on those skills based on abstraction (algebra, geometry, trigonometry) (BNO, 2004; Jiménez-Fernández, 2016). Dyscalculia is not only a disorder in number and mathematical processing, comprehension, and memorization; it includes other disorders in other areas of mathematics, such as visual, spatial, and verbal problems (Menon et al., 2020; Witzel & Mize, 2018). Dyscalculia often co-occurs with dyslexia, but co-occurrence has rarely been studied (Joyner & Wagner, 2020).

2 Some characteristics of student performance gaps and extracurricular activities

Student performance is influenced by a number of factors, such as family background, personal characteristics, motivation, active involvement in lessons, social and territorial environment, and teacher characteristics. The family background plays an important role in many aspects of the child's

development – for example, children from families with a better social background have better health and may be less likely to be born prematurely, which is assumed to make them less likely to develop learning disorders later in life (Parasuraman et al., 2018). Family background also plays an important role in the extent to which parents, siblings, and those in the child's immediate environment can accept the child's difficulties, create a motivating environment for the child, and set an example for the child, for example in communication or schooling (Bernstein, 2003; Hegedűs, 2020; Mesterházi, 2008). Furthermore, it is important to know what kind of school parents send their children to, how involved they are in their children's studies and school life, and whether they can afford to participate in various extracurricular activities (Hegedűs, 2020; Pusztai, 2009; Sebestyén, 2023; Šeďová & Sedláček, 2023; Ule & Živoder, 2023).

According to a previous study, children with learning disorders in Hungary have poorer family backgrounds (Hegedűs, 2021), which is a problem because the literature shows that children with better family backgrounds tend to attend higher quality educational institutions or even choose special "training types"[1] that their parents consider more challenging than the standard curriculum, so overall there is selection in the education system that may be reinforced by institutional admission procedures (Garami, 2020; Hegedűs, 2016, 2020). Significant differences can be seen between students with learning disorders and students with typical development in different training types in general, as well as when comparing the performance of different training types. The lowest performers are students in the standard curriculum and the best performers are students in the training type with extra hours, which includes an increased number of lessons in a subject (Hegedűs & Sebestyén, 2023). Children from better family backgrounds are financially less likely to travel long distances and can finance more extracurricular activities; this can have an overall positive impact on their academic performance (Hegedűs, 2020; Sebestyén, 2023).

Student performance is affected by the type of learning disorder (Hegedűs, 2021), and it is necessary to consider not only – for example – the negative impact of a reading disorder on reading comprehension performance, but also about other areas that may be affected. According to Hegedűs and Sebestyén (2019) reading comprehension, mathematics, and foreign language performance are closely correlated. Depending on the nature of the special

[1] The Act CXC of 2011 on National Public Education gives students the opportunity to study certain subjects for a higher number of lessons. These are known as modular training types or training types with extra hours (Act CXC of 2011).

educational need, children may be co-educated with their mainstream peers, which also affects their academic performance, along with the number of children per special needs teacher in their compulsory development even with co-education (Hegedűs, 2023).

If an individual has a difficulty, such as a learning disorder, it may often be associated with other disorders, such as behavioral problems (Benz & Powell, 2020). This can have a negative impact on the individual's motivation and performance, which may lead to lower goals for future plans (Hegedűs & Sebestyén, 2023; Lohmann, 1998). It is necessary to motivate learners; people with learning disorders may be gifted in other areas, but this is often difficult to see because, according to experts on the relationship between learning disorders and giftedness, there is a potential for masking – children with specific learning disorders may compensate for the emergence of their learning disorder by using their giftedness (Flanagan et al., 2013; Romano et al., 2024; Stankovska & Rusi, 2014).

For children with special educational needs, it is perhaps even more important to address their talents in order to balance their learning difficulties and help them to experience success in other areas (Westendorp et al., 2011). In many cases, giftedness can be addressed through extracurricular activities in or out of school, including support for children from parents and teachers (Beckmann & Minnaert, 2018). The financial situation of the family influences the extracurricular activities that children can attend – from activities organized at school, which are usually free, to private lessons that parents have to pay for (Rolff et al., 2008; Sebestyén, 2023). Research by Szemerszki (2020) showed that children from better family backgrounds are more likely to attend extra lessons in mathematics for catching up and in foreign languages for acquiring additional skills. More educated mothers are more supportive of their children's participation in sports (Şahin, 2018), but this should be important for all, as physical sports activities can contribute to better academic performance (Adeyemo, 2010; Shulruf et al., 2008).

3 Research questions and hypotheses

Based on the literature, we formulated the following research questions for our study:

- What are the differences in reading comprehension and mathematics competence performance scores between students with learning disorders at classes with different curriculum?
- What are the differences between the in-school and extracurricular activities of students with learning disorders at classes with different curriculum?

- What is the relationship between family background index, learning disorder, and classes with different curriculum in performance and participation in-school and extracurricular activities?

In relation to the research questions, the following three hypotheses were formulated:

- Students with learning disorders in the classes with standard curriculum have the lowest mathematics and reading comprehension scores; students without learning disorders in classes with increased numbers of lessons have the highest scores.
- Students without learning disorders participate more in in-school and extracurricular activities in all analyzed class types.
- Family background influences student performance and participation in in-school and extracurricular activities.

4 Database and methods

For this research, we analyze the data from the National Competency Measurement (NCM) 2019, tenth grade,[2] because this is the last database that was recorded before COVID, and the results of the students are not affected by the changed (educational) opportunities during the pandemic. The National Competency Measurement is also called the "little sister of PISA" in Hungary because it measures competence performance in mathematics and reading comprehension. PISA and NCM scores are not comparable because they use a different scoring system. Student competences in mathematics and reading comprehension have been measured using NCM with the participation of all sixth, eighth, and tenth grade students since 2002. A set of centralized tests is prepared for students to complete in their schools on a date in May and under conditions set by the Educational Authority. Background questionnaires are also produced for the competence tests, which include questions about students at the student, site, and institutional level (Oktatási Hivatal, 2024). Our study presents a secondary analysis carried out on the NCM database.

The database contains information on a total of 83,751 students from tenth grade, of whom 1,515 have a learning disorder (dyslexia, dysgraphia, dyscalculia). Learning disorders are diagnosed by special education teachers, psychologists and, if necessary, doctors at educational services. No diagnosis is made during the competence measurement; the students assigned to the examined categories received their diagnosis in advance, and this must be

[2] It is the second year of the ISCED 3.

indicated in the background questionnaires of the competence measurement. The Committee of Educational Experts follows the criteria of the International Statistical Classification of Diseases and Related Health Problems (ICD) for the diagnosis of learning disorders. The different types of learning disorders can thus be interpreted internationally.

Each of the eight types of learning disorders was queried separately in a background questionnaire during the data collection process, so that we were able to identify the 1,515 students with dyslexia, dysgraphia, and/or dyscalculia along these variables. Some students have two or all three of the learning disorders we included in the study, so we took care in the coding to ensure that the learning disorder variables were not simply a combination of the dyslexia, dysgraphia, and dyscalculia groups; students with more than one disorder were also included in the study group. It is possible that a student may have an additional learning disorder type, but this was not investigated because these three types of learning disorders under consideration are the dominant ones, according to the newer classification systems. For the remainder of this article, for the sake of simplicity, the term "learning disorder" is used to refer only to those with dyslexia, dysgraphia, dyscalculia, or some combination of the three, and "learning disorder non-disabled" to refer to those students who do not have dyslexia, dysgraphia, or dyscalculia.

By classes, we can distinguish between students in classes with standard curriculum (37,888), students with a proven language background (11,011) and students in intensive classes (12,332). The NCM database classifies students as standard, bilingual, intensive, and minority language learned. The bilingual and minority types were combined into classes with intensive foreign language learning (intensive FLL), to which we also added students from the standard and the upper-secondary groups who had attended language preparation after eighth grade.[3] The creators of the NCM database classified the students studying in some intensive class into the group with an increased number of hours, but did not indicate which subject they studied for an increased number of hours. Therefore, this group also included students participating in a language course (i.e., students with an increased number of foreign language classes), but the limitation of our research is that we could not separate them

[3] In the Hungarian education system, primary education can last four, six, or eight years. In the first two cases, students can continue their studies in an eight- or six-year grammar school; after eight years of primary education, they can enrol in a four-year grammar school or a three- or four-year vocational school. Secondary school and, in some cases, technical and vocational education and training may be extended by a so-called "zero year" of advanced foreign language learning for bilingual or preparatory language training. This is followed by a second year of secondary school curriculum (Act CXC of 2011; Act LXXX of 2019; Eurydice, 2024).

from those attending other classes. Another limitation of the research is that the background questionnaires for the competence measurement do not record the number of increased lessons. Usually this means two extra lessons, but students have 17 lessons in the language year of dual language courses (Oktatási Hivatal, 2020). For the rest of the standard curriculum and the increased timetable group, we kept the original NCM classification.

To examine family background, we use the central Family Background Index (FBI), a multivariate index created by the Hungarian Educational Authority, which includes parents' educational attainment and workplace. The family background index is composed of the educational attainment of the father and mother, the number of computer(s) in the household, the number of books in the home, and the number of books owned by the student. Higher values indicate higher family background (Hegedűs, 2020). In all the questions where students were asked whether they had participated in development, talent management, etc., yes and no answers were possible; students who did not complete this question were excluded from the assessment.

Secondary analysis of the data is performed using SPSS, two- and three-dimensional ANOVA analysis, and two- and three-dimensional cross-tabulation analysis using a Chi-square test and row percentage. In the family background index, it was possible to use the variables combined in the index as one variable during the ANOVA analysis, rather than binary coded in the regression analysis. In the analysis, asterisks are used to indicate significant relationships, and the number of asterisks indicates the strength of significance (***$p \leq 0.001$, **$0.001 < p \leq 0.010$, *$0.010 < p \leq 0.050$). A limitation of our research is that not all students have all their data filled in, so there may be differences in the number of students during the analyses presented in the study.

5 Results

5.1 Student performance according to learning disorder and class types

Significant connections were found according to learning disorder and training type between mathematics, reading comprehension, and family background index. The lowest scores were found in the standard curriculum classes, followed by those classes with intensive FLL; while the highest scores were found in those classes with an increased number of lessons, regardless of whether the student had a learning disorder or not. In all grades, the results of students with learning disorders were significantly lower than those of students without learning disorders.

We also found a difference in the family background index, because students with learning disorders had a poorer background in the classes with an increased number of lessons, as well as in the classes with intensive FLL.

In contrast, the family background of students with learning disorders was minimally better in the classes with an increased number of lessons. The selection by family background could also be identified by the different choices of training types (Table 1).

Table 1
Correlations between NCM scores and family background index of students with and without learning disorders by training types

Training type	Learning disorder		Mathematics***	Reading***	Family background index***
with standard curriculum	no	*M*	1619.0	1607.0	–0.208
		SD	194.3	195.7	0.937
	yes	*M*	1504.7	1468.9	–0.257
		SD	178.6	175.9	0.821
	total	*M*	1616.2	1603.6	–0.209
		SD	194.7	196.4	0.934
intensive FLL	no	*M*	1718.3	1717.9	0.303
		SD	178.9	183.9	0.865
	yes	*M*	1561.2	1526.1	0.100
		SD	194.0	196.9	0.825
	total	*M*	1716.9	1716.2	0.301
		SD	179.6	184.9	0.865
with an increased number of lessons	no	*M*	1781.6	1778.5	0.524
		SD	157.8	148.0	0.743
	yes	*M*	1636.4	1623.7	0.534
		SD	209.2	182.6	0.727
	total	*M*	1780.6	1777.4	0.524
		SD	158.7	148.9	0.743
total	no	*M*	1670.2	1662.1	0.035
		SD	196.6	198.4	0.941
	yes	*M*	1520.3	1486.4	–0.160
		SD	186.3	183.5	0.846
	total	*M*	1667.5	1658.9	0.032
		SD	197.4	199.5	0.940
N			60,534	60,572	56,640

Note: ***$p \leq 0.001$ Source: NCM (2019)

5.2 In-school and extracurricular activities of students with and without learning disorders in different classes

In examining the activities within the institution, we conducted a multi-dimensional cross-tabulation analysis with a row percentage. The database

did not allow us to examine whether students with learning disorders participate in the educationally required habilitation and rehabilitation activities, but Table 2 shows that students without learning disorders might also participate in similar remedial activities. Participation in remediation and development activities showed an inverted pattern for students with and without learning disorders. Of the students without a learning disorder, 18.5% of those in the standard curriculum participated in such activities, 17.2% of those from classes with intensive FLL, and 16.1% of those from classes with an increased number of lessons; for students with a learning disorder, 60.3% were in the standard curriculum, 69.0% in the intensive FLL, and 76.5% in the classes with an increased number of lessons.

Table 2
Participation of students with and without learning disorders in tutoring and developmental activities by training types

Training type	Learning disorder		Participation in tutoring, developmental activities		Total
			yes	no	
with standard curriculum***	no	*N*	6,301	27,720	34,021
		row%	18.5%	81.5%	100.0%
	yes	*N*	511	337	848
		row%	60.3%	39.7%	100.0%
	total	*N*	6,812	28,057	34,869
		row%	19.5%	80.5%	100.0%
intensive FLL***	no	*N*	1,714	8,255	9,969
		row%	17.2%	82.8%	100.0%
	yes	*N*	60	27	87
		row%	69.0%	31.0%	100.0%
	total	*N*	1,774	8,282	10,056
		row%	17.6%	82.4%	100.0%
with an increased number of lessons***	no	*N*	1,811	9,457	11,268
		row%	16.1%	83.9%	100.0%
	yes	*N*	65	20	85
		row%	76.5%	23.5%	100.0%
	total	*N*	1,876	9,477	11,353
		row%	16.5%	83.5%	100.0%
total	no	*N*	9,826	45,432	55,258
		row%	17.8%	82.2%	100.0%
	yes	*N*	636	384	1,020
		row%	62.4%	37.6%	100.0%
	total	*N*	10,462	45,816	56,278
		row%	18.6%	81.4%	100.0%

Note: ***$p \leq 0.001$ Source: NCM (2019)

Half as many students attended talent development education as attended tutoring or catch-up courses. There was a significant difference ($p = 0.033$) between the two groups in the classes with standard curriculum, in which the proportion of students with learning disorders attending talent development education was the lowest of all groups (5.5%). Another significant difference ($p = 0.001$) was found for the whole sample of the database in favor of those without a learning disorder. A higher proportion of students participated in talent development education than in the intensive FLL, where the participation rate of students with learning disorders (11.4%) was higher than that of students without learning disorders (10.6%). The highest participation rate in talent development education was found in the classes with an increased number of lessons, with 15.6% of students with learning disorders participating in talent development education compared to 16.9% of students without learning disorders. They also showed that different classes offer different opportunities for students (Table 3). We also ran a similar analysis of the sports opportunities provided by the institution, but in neither case did the results show a significant difference.

Table 3
Participation of students with and without learning disorders in gifted education by training type

Training type	Learning disorder		Participation in a talent development education		Total
			yes	no	
with standard curriculum*	no	*N*	2,437	30,944	33,381
		row%	7.3%	92.7%	100.0%
	yes	*N*	43	738	781
		row%	5.5%	94.5%	100.0%
	total	*N*	2,480	31,682	34,162
		row%	7.3%	92.7%	100.0%
intensive FLL	no	*N*	1,041	8,787	9,828
		row%	10.6%	89.4%	100.0%
	yes	*N*	9	70	79
		row%	11.4%	88.6%	100.0%
	total	*N*	1,050	8,857	9,907
		row%	10.6%	89.4%	100.0%
with an increased number of lessons	no	*N*	1,898	9,315	11,213
		row%	16.9%	83.1%	100.0%
	yes	*N*	12	65	77
		row%	15.6%	84.4%	100.0%
	total	*N*	1,910	9,380	11,290
		row%	16.9%	83.1%	100.0%
total**	no	*N*	5,376	49,046	54,422
		row%	9.9%	90.1%	100.0%
	yes	*N*	64	873	937
		row%	6.8%	93.2%	100.0%
	total	*N*	5,440	49,919	55,359
		row%	9.8%	90.2%	100.0%

Note: **$0.001 < p \leq 0.010$, *$0.010 < p \leq 0.050$ Source: NCM (2019)

The following was a description of activities outside the institution. There was a significant difference ($p = 0.004$) between the two groups in the participation in extra math lessons at the classes with an increased number of lessons. At this level, 23.2% of students without learning disorders participated in math lessons compared to 36.3% of students with learning disorders. There was also a difference between the training types, with the lowest participation in math lessons in the standard curriculum (16.4% and 14.8%) and higher in the intensive FLL (18.5% and 17.6%). In comparison, participation was much higher in the classes with an increased number of lessons (Table 4). It should be noted that the background questionnaire of the competence measurement does not cover Hungarian language lessons, so we were not able to investigate participation in these classes.

Table 4
Participation of students with and without learning disorders in special mathematics lessons by training types

Training type	Learning Disorder		Participation in a special mathematics lesson		Total
			yes	no	
with standard curriculum	no	*N*	5,589	28,574	34,163
		row%	16.4%	83.6%	100.0%
	yes	*N*	123	710	833
		row%	14.8%	85.2%	100.0%
	total	*N*	5,712	29,284	34,996
		row%	16.3%	83.7%	100.0%
intensive FLL	no	*N*	1873	8,225	10,098
		row%	18.5%	81.5%	100.0%
	yes	*N*	15	70	85
		row%	17.6%	82.4%	100.0%
	total	*N*	1,888	8,295	10,183
		row%	18.5%	81.5%	100.0%
with an increased number of lessons*	no	*N*	2,631	8,708	11,339
		row%	23.2%	76.8%	100.0%
	yes	*N*	29	51	80
		row%	36.3%	63.8%	100.0%
	total	*N*	2,660	8,759	11,419
		row%	23.3%	76.7%	100.0%
total	no	*N*	10,093	45,507	55,600
		row%	18.2%	81.8%	100.0%
	yes	*N*	167	831	998
		row%	16.7%	83.3%	100.0%
	total	*N*	10,260	46,338	56,598
		row%	18.1%	81.9%	100.0%

Note: *$0.010 < p \leq 0.050$ Source: NCM (2019)

There were two significant differences in participation in extracurricular foreign language classes. In the standard curriculum ($p = 0.027$), 14.9% of students without learning disorders attended such extracurricular classes outside the institution, compared to 12.5% of students with learning disorders. The other significant result ($p = 0.000$) was in the full dataset, where 20.1% of students without learning disorders attended extracurricular classes in extracurricular foreign languages, compared to 15.1% of students with learning disorders. There was no significant difference between the two groups for those classes with intensive FLL, but a much higher proportion of students attended special foreign language classes (21.9% and 19.8%), with the highest rates for students in classes with an increased number of lessons, where more than a third of

students (34.1% and 37.5%) attended extracurricular foreign language classes. In the latter case, a higher proportion of students with learning disorders attended an extracurricular foreign language class, which was not the case in other training types. The higher proportions in the classes with an increased number of lessons might be explained not only by the better family background of the students (see Table 1), but also by the unknown proportion of students from classes with an increased number of foreign language lessons (Table 5).

Table 5
Participation of students with and without learning disorders in extracurricular foreign language lessons by training types

Training type	Learning disorder		Participation in extracurricular foreign language lessons		Total
			yes	no	
with standard curriculum*	no	*N*	5,067	28,862	33,929
		row%	14.9%	85.1%	100.0%
	yes	*N*	102	717	819
		row%	12.5%	87.5%	100.0%
	total	*N*	5,169	29,579	34,748
		row%	14.9%	85.1%	100.0%
intensive FLL	no	*N*	2,185	7,808	9,993
		row%	21.9%	78.1%	100.0%
	yes	*N*	16	65	81
		row%	19.8%	80.2%	100.0%
	total	*N*	2,201	7,873	10,074
		row%	21.8%	78.2%	100.0%
with an increased number of lessons	no	*N*	3,876	7,489	11,365
		row%	34.1%	65.9%	100.0%
	yes	*N*	30	50	80
		row%	37.5%	62.5%	100.0%
	total	*N*	3,906	7,539	11,445
		row%	34.1%	65.9%	100.0%
total*	no	*N*	11,128	44,159	55,287
		row%	20.1%	79.9%	100.0%
	yes	*N*	148	832	980
		row%	15.1%	84.9%	100.0%
	total	*N*	11,276	44,991	56,267
		row%	20.0%	80.0%	100.0%

Note: *$0.010 < p \leq 0.050$ Source: NCM (2019)

Extra sports lessons could also have a beneficial effect on student performance, so we included this in the study. In the full dataset, there was a significant difference ($p = 0.000$) between the two groups of students, with a higher proportion of students without learning disorders (28.6%) attending extra sport lessons than students with learning disorders (23.7%). At the classes with standard curriculum, close to the borderline of significance ($p = 0.053$), we saw that students without learning disorders have a higher proportion of participation in extra sports lessons (23.7% and 21.2%). The proportion of students with intensive FLL who participated in extra sports classes is much higher compared to the standard curriculum classes (34.9% and 34.1%); the proportion was even higher in classes with an increased number of lessons (37.9% and 38.3%) (Table 6).

Table 6
Participation of students with and without learning disorders in extra sports lessons by training types

Training type	Learning disorder		Participation in extra sports lessons		Total
			yes	no	
with standard curriculum	no	*N*	8,025	25,885	33,910
		row%	23.7%	76.3%	100.0%
	yes	*N*	173	644	817
		row%	21.2%	78.8%	100.0%
	total	*N*	8,198	26,529	34,727
		row%	23.6%	76.4%	100.0%
intensive FLL	no	*N*	3,464	6,457	9,921
		row%	34.9%	65.1%	100.0%
	yes	*N*	28	54	82
		row%	34.1%	65.9%	100.0%
	total	*N*	3,492	6,511	10,003
		row%	34.9%	65.1%	100.0%
with an increased number of lessons	no	*N*	4,256	6,960	11,216
		row%	37.9%	62.1%	100.0%
	yes	*N*	31	50	81
		row%	38.3%	61.7%	100.0%
	total	*N*	4,287	7,010	11,297
		row%	37.9%	62.1%	100.0%
total***	no	*N*	15,745	39,302	55,047
		row%	28.6%	71.4%	100.0%
	yes	*N*	232	748	980
		row%	23.7%	76.3%	100.0%
	total	*N*	15,977	40,050	56,027
		row%	28.5%	71.5%	100.0%

***$p \leq 0.001$ Source: NCM (2019)

As shown in Table 7, we examined the family background index of students who participated and did not participate in in-school activities as a function of training type and learning disorder, in order to obtain information on the role of family background in participation in extracurricular activities. We found a significant difference in all categories by ANOVA analysis ($p = 0.000$). In all three training types, students without learning disorders who did not participate in tutoring had a better family background index. Students with a learning disorder were more likely to participate in these activities if they had a better family background; the reverse was true for those with a learning disorder, but the values were very high and close together.

In terms of participation in talent development education, there was a clear trend in all three training types, in both groups of students, towards participation by students with a better family background, which suggests that family background played a decisive role. Our research showed that participation in school recreational sports activities per se did not show significant differences between the groups of students studied, but that differences could be found when family background was examined. There were small differences in the family background index for participation in school sports activities for students in the classes with standard curriculum, while in the intensive FLL and classes with an increased number of lessons there was a tendency for those with a higher family background to participate less in these activities (Table 7).

Table 7
Family backgrounds of students with and without learning disorders in different extracurricular school activities by training types

Training type	Learning disorder	Tutoring, development lessons	FBI***	Talent management lessons	FBI ***	School recreational sports activities	FBI***
with standard curriculum	no	yes	–0.213	yes	0.378	yes	–0.201
		no	–0.205	no	–0.248	no	–0.207
		total	–0.207	total	–0.202	total	–0.206
	yes	yes	–0.176	yes	0.293	yes	–0.269
		no	–0.359	no	–0.270	no	–0.252
		total	–0.248	total	–0.240	total	–0.255
	total	yes	–0.210	yes	0.377	yes	–0.202
		no	–0.207	no	–0.248	no	–0.208
		total	–0.208	total	–0.203	total	–0.207
intensive FLL	no	yes	0.209	yes	0.443	yes	0.233
		no	0.331	no	0.295	no	0.328
		total	0.310	total	0.310	total	0.310
	yes	yes	0.137	yes	0.285	yes	–0.061
		no	–0.218	no	0.097	no	0.128
		total	0.034	total	0.113	total	0.092
	total	yes	0.207	yes	0.442	yes	0.231
		no	0.329	no	0.293	no	0.327
		total	0.307	total	0.309	total	0.308
with an increased number of lessons	no	yes	0.465	yes	0.754	yes	0.539
		no	0.538	no	0.483	no	0.520
		total	0.526	total	0.529	total	0.523
	yes	yes	0.525	yes	0.666	yes	0.336
		no	0.560	no	0.512	no	0.577
		total	0.533	total	0.536	total	0.545
	total	yes	0.467	yes	0.753	yes	0.538
		no	0.538	no	0.484	no	0.520
		total	0.526	total	0.529	total	0.523
total	no	yes	–0.012	yes	0.524	yes	0.035
		no	0.049	no	–0.010	no	0.040
		total	0.038	total	0.043	total	0.039
	yes	yes	–0.072	yes	0.372	yes	–0.208
		no	–0.299	no	–0.180	no	–0.146
		total	–0.156	total	–0.144	total	–0.156
	total	yes	–0.016	yes	0.523	yes	0.031
		no	0.046	no	–0.013	no	0.037
		total	0.035	total	0.040	total	0.036
N			52,564	N	51,741	N	51,616

***$p \leq 0.001$ Source: NCM (2019)

5.3 Family background index according to learning disorder, class types, and participation in-school and extracurricular activities

In our final analysis, shown in Table 8, we examined extracurricular activities in the context of learning disorder, training type, and family background index. For participation in extra lessons in mathematics, the effect of family background was clearly visible, as students with a better family background were more likely to participate in extra lessons, which could also be explained by the fact that most of these lessons had to be paid for. There was one exception, in the category of students with learning disorders in classes with an increased number of lessons, where the family background index of students attending extra lessons was lower (0.517 and 0.530).

A similar trend could be observed in the case of extra tuition in a foreign language, where those with a better family background were more likely to attend this type of extra tuition. On the other hand, it could be seen that those from lower family backgrounds attended mathematics classes; those from better family backgrounds spent more on foreign language classes.

The extra sports classes were also attended by students from better family backgrounds, and the family background index increased in the direction of the classes with an increased number of lessons. In general, it was true for the participation in sport classes that the family background index was between the family background index in each category between the extra lessons in mathematics and foreign language, with exceptions such as students with learning disorders in the classes with standard curriculum or students without learning disorders in the classes with an increased number of lessons (Table 8).

Table 8
Family background of students with and without learning disorders in different extracurricular activities by training type

Training type	Learning disorder	Mathe-matics	FBI***	Foreign language	FBI***	Sport	FBI***
with standard curriculum	no	yes	0.074	yes	0.426	yes	0.226
		no	–0.260	no	–0.315	no	–0.339
		total	–0.205	total	–0.203	total	–0.204
	yes	yes	0.063	yes	0.207	yes	–0.024
		no	–0.328	no	–0.324	no	–0.318
		total	–0.269	total	–0.257	total	–0.257
	total	yes	0.074	yes	0.422	yes	0.221
		no	–0.262	no	–0.316	no	–0.338
		total	–0.207	total	–0.204	total	–0.205
intensive FLL	no	yes	0.392	yes	0.535	yes	0.582
		no	0.287	no	0249	no	0.160
		total	0.307	total	0.312	total	0.308
	yes	yes	0.503	yes	0.311	yes	0.469
		no	0.052	no	0.080	no	0.081
		total	0.132	total	0.128	total	–0.119
	total	yes	0.393	yes	0.534	yes	0.581
		no	0.285	no	0.248	no	0.158
		total	0.305	total	0.311	total	0.307
with an increased number of lessons	no	yes	0.554	yes	0.671	yes	0.739
		no	0.516	no	0.448	no	0.392
		total	0.525	total	0.525	total	0.524
	yes	yes	0.517	yes	0.798	yes	0.665
		no	0.530	no	0.392	no	0.478
		total	0.525	total	0.538	total	0.551
	total	yes	0.554	yes	0.672	yes	0.738
		no	0.516	no	0.448	no	0.393
		total	0.525	total	0.525	total	0.524
total	no	yes	0.259	yes	0.533	yes	0.444
		no	–0.011	no	–0.084	no	–0.126
		total	0.039	total	0.042	total	0.038
	yes	yes	0.182	yes	0.339	yes	0.140
		no	–0.241	no	–0.247	no	–0.246
		total	–0.169	total	–0.157	total	–0.155
	total	yes	0.258	yes	0.530	yes	0.440
		no	–0.015	no	–0.087	no	–0.128
		total	0.035	total	0.038	total	0.035
N			52,838	N	52,563	N	52,320

***$p \leq 0.001$

Source: NCM (2019)

6 Discussion

In testing our hypotheses, we found that our first hypothesis – students with learning disorders in the classes with standard curriculum have the lowest mathematics and reading comprehension scores, while students without learning disorders in the classes with an increased number of lessons have the highest scores – was confirmed. The largest difference between the categories was 277 points in mathematics and nearly 310 points in reading comprehension, which could be considered very significant.

Our second hypothesis – students without learning disorders would participate more in in-school and extracurricular activities in all analyzed class types – was partially confirmed. A higher proportion of students with a learning disorder participated in tutoring and catch-up activities, but this was understandable given the nature of the learning disorder. The results showed that a higher proportion of students without learning disorders participate in the classes with standard curriculum and in intensive FLL, both inside and outside school; this was not the case for the classes with an increased number of lessons.

Our third hypothesis – family background influences student performance and participation in in-school and extracurricular activities – was confirmed. Students with better family backgrounds achieved higher results than both groups of students tested. By training types, there were much larger differences in family background in favor of the classes with an increased number of lessons. Family background was a determinant of student attendance in – presumably paid – extra lessons, for which the data suggest that parents appeared to have priority.

7 Conclusion

The aim of the study was to investigate the relationship between the training type, the presence of a learning disorder, participation in in-school and extracurricular activities, and the results of the competency measurement. In this study, we conducted a secondary analysis on the dataset of the National Competency Measurement of all tenth grade students in Hungary in 2019. The study of students with learning disorders is important because it is the largest group within the special educational needs category and the largest majority of these students who are integrated learners. In Hungary, educational selection on the basis of family background is very significant; this is also reflected in the differences between the different classes. The data show that those from the best family backgrounds are in the classes with an increased number of lessons, while those from the poorest family backgrounds are in

the classes with standard curriculum. Family background also determines the proportion of students who participate in activities inside and outside school. Priorities can also be seen in that even students from lower family backgrounds participate in extra lessons in mathematics, while less so in extra lessons in foreign languages or sports. There are also marked differences between the learning-disabled and non-learning-disabled groups by training type, with students with learning disorders scoring lower on competence performance outcomes. Students with a learning disorder have a much higher rate of participation in tutoring or catch-up courses. Family background, class, and participation in extracurricular activities are closely linked, because the better the family background, the more likely the student is to be in classes with an increased number of lessons, where the participation rate in extracurricular activities is much higher.

In the long term, the school system should provide care for students with learning disorders that would reduce disadvantages resulting from their family background. The school system should place more emphasis on students with learning disorders, which would require a central change. The research could be taken further by looking at concrete examples and at teacher views of the development of these students.

References

Act CXC of 2011 on National public education (2011). https://net.jogtar.hu/jogszabaly?docid=a1100190.tv

Act LXXX of 2019 on Vocational education and training (2014). https://net.jogtar.hu/jogszabaly?docid=a1900080.tv

Adeyemo, S. A. (2010). The relationship between students' participation in school based extracurricular activities and their achievement in physics. *International Journal of Science and Technology Education Research, 1*(6), 111–117.

Beckmann, E., & Minnaert, A. (2018). Non-cognitive characteristics of gifted students with learning disabilities: An in-depth systematic review. *Frontiers in Psychology, 9*, 504. https://doi.org/10.3389/fpsyg.2018.00504

Benz, S. A., & Powell, S. R. (2020). The influence of behavior on performance within a word-problem intervention for students with mathematics difficulty. *Remedial and Special Education, 42*(3), 182–192. https://doi.org/10.1177/0741932520923063

Bernstein, B. B. (2003). *Class, codes and control*. Routledge.

Biotteau, M., Danna, J., Puyjarinet, F., Velay, J.-L., Albaret, J.-M., & Chaix, Y. (2019). Developmental coordination disorder and dysgraphia: Signs and symptoms, diagnosis, and rehabilitation. *Neuropsychiatric Disease and Treatment, 15*, 1873–1885. https://doi.org/10.2147/ndt.s120514

BNO. (2004). *Betegségek Nemzetközi Osztályozása* [International classification of diseases]. BNO.

Chase, C., & Tallal, P. (1992). Learning disabilities: Cognitive aspects. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 219−225). Macmillan Publishing Company.

Chordia, S. L., Thandapani, K., & Arunagirinathan, A. (2020). Children 'at risk' of developing specific learning disability in primary schools. *Indian Journal of Pediatrics, 87*(2), 94–98. https://doi.org/10.1007/s12098-019-03130-z

Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and dysgraphia: definition, diagnosis, and management. *Translational Pediatrics, 9*(1), 46–54. http://dx.doi.org/10.21037/tp.2019.11.01

Cristofani, P., Di Lieto, M. C., Casalini, C., Pecini, C., Baroncini, M., Pessina, O., Gasperini, F., Dasso Lang, M. B., Bartoli, M., Chilosi, A. M., & Milone, A. (2023). Specific learning disabilities and emotional-behavioral difficulties: Phenotypes and role of the cognitive profile. *Journal of Clinical Medicine, 12*(5), 1882. https://doi.org/10.3390%2Fjcm12051882

Csépe, V. (2008). A különleges oktatást, nevelést és rehabilitációs célú fejlesztést igénylő (SNI) gyermekek ellátásának gyakorlata és a szükséges teendők [The practice of care for children requiring special education, upbringing and development for the purpose of rehabilitation (SEN) and the necessary actions]. In K. Fazekas, J. Köllő, & J. Varga (Eds.), *Zöld könyv a magyar közoktatás megújításáért 2008* (pp. 139–165). Oktatás és Gyermekesély Kerekasztal, Ecostat.

Dékány, J., & Mohai, K. (2012). Egyéb pszichés fejlődési zavarral küzdő gyermekek, tanulók komplex vizsgálatának diagnosztikus protokollja – Specifikus tanulási zavarok (írott nyelvhasználat zavarai, diszkalkulia) [Diagnostic protocol for the complex examination of children and students with other mental developmental disorders – Specific learning disorders (written language disorders, dyscalculia)]. In Á. Torda (Ed.), *Diagnosztikai kézikönyv*. Educatio Társadalmi Szolgáltató Nonprofit Kft. https://paks.tmpsz.hu/userfiles/files/diagnosztikai_kezikonyv_9fejezet.pdf

Englbrecht, A., & Weigert, H. (1996). *Hogyan akadályozzuk meg a tanulási akadályok kialakulását? avagy Nem jelenthet akadályt a tanulási akadály!* [How do we prevent the development of learning barriers? or The learning barriers cannot be an obstacle!]. Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola.

Eurydice. (2024, 17th July). *Hungary: Overview*. Eurydice. https://eurydice.eacea.ec.europa.eu/national-education-systems/hungary/overview

Farkasné Gönczi, R. (2011). A diszkalkulia – Terminológiaváltás, a diagnosztika forradalma terápia megújulása [Dyscalculia – Change of terminology, revolution of diagnostics, renewal of therapy]. In G. Papp (Ed.), *A diagnózistól a foglalkozási rehabilitációig: új utak és eredmények a tanulásban akadályozott személyek gyógypedagógiájában* (pp. 197–224). ELTE Bárczi Gusztáv Gyógypedagógiai Kar, Eötvös Kiadó.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd Ed.). John Wiley and Sons.

Garami, E. (2020). A területi környezet egyenlőtlenségei [Inequalities of the territorial environment]. In G. Pusztai (Ed.), *Nevelésszociológia. Elméletek, közösségek, kontextusok* (pp. 348–360). Debreceni Egyetemi Kiadó.

Gerebenné Várbíró, K. (1995). A tanulási zavar jelenségkörének gyógypedagógiai pszichológiai értelmezése [Psychological interpretation of the phenomenon of learning disorder in special education]. In P. Zászkaliczky (Ed.), *„…önmagában véve senki sem …„. Tanulmányok a gyógypedagógiai pszichológia és a határtudományok köréből. Lányiné dr. Engelmayer Ágnes 65. születésnapjára* (pp. 216–245). Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola.

Gyarmathy, É. (1998). Tanulási zavarok azonosítása és kezelése az óvodában és az általános iskolában [Identification and treatment of learning disorders in kindergarten and elementary school]. *Új Pedagógiai Szemle, 48*(11), 68–76.

Gyarmathy, É. (2018). A nyelvi fejlődés nehézségei és a diszlexia [Difficulties in language development and dyslexia]. *Gyermeknevelés Tudományos Folyóirat, 6*(3), 77–92. https://doi.org/10.31074/201837792

Hegedűs, R. (2016). Tizedik osztályos tanulók teljesítményének területi különbségei [Territorial differences in performance of 10th grade students]. *Iskolakultúra, 26*(12), 16–30. https://doi.org/10.17543/ISKKULT.2016.12.16

Hegedűs, R. (2020). *Kompetenciák – hátrányok – térségek. Avagy honnan s hogyan jutnak el a hátrányos helyzetűek a felsőoktatásba?* [Competences – disadvantages – territorials. From where and how can get disadvantaged students into the higher education system?]. Debreceni Egyetemi Kiadó.

Hegedűs, R. (2021). A tanulási zavarral küzdő gyermekek Országos kompetenciamérésen elért eredményei és jellemzői [The results and characteristics of children with learning disorders on the National Competence Measurement]. *Gyógypedagógiai Szemle, 69*(4), 262–277. https://doi.org/10.52092/gyosze.2021.4.2

Hegedűs, R. (2023). Sajátos nevelési igényű gyermekek, tanulók számának területi és statisztikai elemzése [Territorial and statistical analysis of the number of children and students with special educational needs]. *Educatio, 32*(2), 228–246. https://doi.org/10.1556/2063.32.2023.2.4

Hegedűs, R. (2024). *Tanulási zavarok és iskolai teljesítmény* [Learning disorders and school performance]. Belvedere Meridionale.

Hegedűs, R., & Sebestyén, K. (2019). A case study about regional differences of Hungarian pupils' achievement. *Hungarian Educational Research Journal, 9*(1), 123–127. https://doi.org/10.1556/063.9.2019.1.12

Hegedűs, R., & Sebestyén, K. (2023). Focus on 10th grade students with learning problems: What influences their achievements? *Hungarian Educational Research Journal, 13*(3), 385–404. https://doi.org/10.1556/063.2022.00168

Jiménez-Fernández, G. (2016). How can I help my students with learning disabilities in mathematics? *REDIMAT – Journal of Research in Mathematics Education, 5*(1), 56–73. https://doi.org/10.17583/redimat.2016.1469

Joyner, R. E., & Wagner, R. K. (2020). Co-occurrence of reading disabilities and math disabilities: A meta-analysis. *Scientific Studies of Reading, 24*(1), 14–22. https://doi.org/10.1080/10888438.2019.1593420

Kirk, S. A., & Bateman, B. (1962). Diagnosis and remediation of learning disabilities. *Exceptional Children, 29*(2), 73–78. https://doi.org/10.1177/001440296202900204

LeRoy, B. W., Samuel, P., Deluca, M., & Evans, P. (2019). Students with special educational needs within PISA. *Assessment in Education: Principles, Policy & Practice, 26*(4), 386–396. https://doi.org/10.1080/0969594X.2017.1421523

Lohmann, B. (1998). *Diszlexiások az iskolában* [Dyslexics at school]. Akkord Kiadó.

Márkus, A. (2007). *Számok, számolás, számolászavarok* [Numbers, counting, counting disorders]. Pro Die Kiadó.

Meixner, I. (2015). *A dyslexia prevenció, reedukáció módszere* [Method of dyslexia prevention and reeducation]. Meixner Alapítvány.

Menon, V., Padmanabhan, A., & Schwartz, F. (2020). Cognitive neuroscience of dyscalculia and math learning disabilities. In K. C. Kadosh (Ed.), *The Oxford handbook of developmental cognitive neuroscience* (pp. 729−764). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198827474.013.23

Mesterházi, Z. (Ed.). (2008). *Inkluzív nevelés. Kézikönyv a szakértői bizottságok működéséhez* [Inclusive education. Manual for the operation of expert committees]. Educatio Társadalmi Szolgáltató Kht.

Mesterházi, Z., & Szekeres, Á. (Eds.). (2019). *A nehezen tanuló gyermekek iskolai nevelése* [School education of children with learning difficulties]. ELTE Bárczi Gusztáv Gyógypedagógiai Kar.

Oktatási Hivatal. (2020). *A 2020-as Nathoz illeszkedő tartalmi szabályozók* [Content regulators to Nat 2020]. Oktatási hivatal. https://www.oktatas.hu/kozneveles/kerettantervek/2020_nat

Oktatási Fivatal. (2024). *Országos kompetenciamérés általános leírás* [National competency measurement general description]. https://www.oktatas.hu/kozneveles/meresek/kompetenciameres/alt_leiras

Parasuraman, S. R., Anglin, T. M., McLellan, S. E., Riley, C., & Mann, M. Y. (2018). Health care utilization and unmet need among youth with special health care needs. *Journal of Adolescent Health, 63*(4), 435–444. https://doi.org/10.1016/j.jadohealth.2018.03.020

Pusztai, G. (2009). *A társadalmi tőke és az iskola. Kapcsolati erőforrások hatása az iskolai pályafutásra* [Social capital and school. The influence of relational resources on the school career]. Új Mandátum Könyvkiadó.

Rolff, H.-G., Leucht, M., & Rösner, E. (2008). Sozialer und familialer Hintergrund [Social and family background]. In DESI-Konsortium (Eds.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (pp. 283–300). Beltz Verlag.

Romano, S., Esposito, D., Aricò, M., Arigliani, E., Cavalli, G., Vigliante, M., Penge, R., Sogos, C., Pisani, F., & Romani, M. (2024). Giftedness and twice-exceptionality in children suspected of ADHD or specific learning disorders: A retrospective study. *Sci, 6*(2), 23. https://doi.org/10.3390/sci6020023

Rourke, B. P. (1993). Arithmetic disabilities, specific and otherwise: A neuropsychological perspective. *Journal of Learning Disabilities, 26*(4), 214–226. https://doi.org/10.1177/002221949302600402

Róbert, P. (2004). Iskolai teljesítmény és társadalmi háttér nemzetközi összehasonlításban [School performance and social background in international comparison]. In T. Kolosi, I. G. Tóth, & G. Vukovich (Eds.), *Társadalmi riport 2004*. (pp. 193–205). TÁRKI.

Şahin, M. (2018). Examination of the level of participation of secondary school students in extracurricular sports activities [Burdur example]. *Asian Journal of Education and Training, 4*(3), 176–181.
https://doi.org/10.20448/journal.522.2018.43.176.181

Şahin, S. K., Kaya Kara, Ö., Köse, B., & Kara, K. (2020). Investigation on participation, supports and barriers of children with specific learning disabilities. *Research in Developmental Disabilities, 101*(12), 103639.
https://doi.org/10.1016/j.ridd.2020.103639

Sebestyén, K. (2023). *Kik és miért tanulnak németet az érettségit adó képzésekben?* [Who and why do they study German in the programs with secondary school leaving exam?]. Debreceni Egyetemi Kiadó.

Shulruf, B., Tumen, S., & Tolley, H. (2008). Extracurricular activities in school, do they matter? *Children and Youth Services Review, 30*(4), 418–426.
https://doi.org/10.1016/j.childyouth.2007.10.012

Stankovska, G., & Rusi, M. (2014). Cognitive, emotional and social characteristics of gifted students with learning disability. In N. Popov (Ed.), *Education's role in preparing globally competent citizens. BCES conference books. Volume 12.* (pp. 438–442). Bulgarian Comparative Education Society.

Szemerszki, M. (2020). Különórák az iskolában és iskolán kívül [Extracurricular lessons in and outside of school]. *Educatio, 29*(2), 205–221.
https://doi.org/10.1556/2063.29.2020.2.3

Šeďová, K., & Sedláček, M. (2023). How vocal and silent forms of participation in combination relate to student achievement. *Instructional Science, 51*(1), 343–361.
https://dx.doi.org/10.1007/s11251-022-09609-1

Ule, M., & Živoder, A. (2023). "Good" parenting: Parental support in education as a factor in inequality. *Studia Paedagogica, 28*(2), 11–28.
https://doi.org/10.5817/SP2023-2-1

Westendorp, M., Hartman, E., Houwen, S., Smith, J., & Visscher, C. (2011). The relationship between gross motor skills and academic achievement in children with learning disabilities. *Research in Developmental Disabilities, 32*(6), 2773–2779.
https://doi.org/10.1016/j.ridd.2011.05.032

Witzel, B., & Mize, M. (2018). Meeting the needs of students with dyslexia and dyscalculia. *SRATE Journal, 27*(1), 31–39.

Žabkar, J., Urankar, T., Javornik, K., & Košak Babuder, M. (2023). Identifying reading fluency in pupils with and without dyslexia using a machine learning model on texts assessed with a readability application. *CEPS Journal, 13*(4), 233–256.
https://doi.org/10.26529/cepsj.1367

STUDIA PAEDAGOGICA

vol. 29, n. 3, 2024
https://doi.org/10.5817/SP2024-3-5

STUDY

INNOVATION IN THE FINAL ASSESSMENT OF STUDENT LEARNING OUTCOMES (REPORT CARDS): THE PROCESS OF CHANGE AND ITS IMPACTS ON SCHOOL PRACTICE

Julie Grombířová[a], Jana Kratochvílová[a]

[a] Faculty of Education, Masaryk University

ABSTRACT

In this paper, we examine the development of a change process in implementing a new approach to the final assessment of student learning outcomes (report cards) in the Czech Republic. The study investigates the development and impacts of the implementation process in primary and lower secondary school. The qualitative research investigation, conducted through repeated interviews with 23 teachers and school leadership, showed that after a year and a half of working with the innovative report card, it had been accepted by teachers in primary school and had significantly influenced the culture of teaching and learning. There was a gradual transformation of the student assessment and self-assessment system and a more profound interconnection between formative and summative assessment. For now, lower secondary school teachers have rejected the innovative report card, although the culture of teaching and learning has been positively influenced.

KEYWORDS

report card; process of change; formative assessment; summative assessment; learning outcomes

CORRESPONDING AUTHOR
Julie Grombířová, Faculty of Education, Masaryk University, Poříčí 7/9, 639 00 Brno, Czech Republic
e-mail: grombirova@ped.muni.cz

Introduction

Systematic assessment of student learning and outcomes has long been an important topic addressed by experts in international and national contexts (e.g. Black & Wiliam, 1998; Earl, 2003; Guskey & Bailey, 2001; Kratochvílová, 2011, 2012, 2015a, 2015b; Nitko & Brookhart, 2014; Slavík, 1999; Starý & Laufková, 2016; Wiliam, 2011a). Many factors influence the demands for changes in assessment: curriculum transformation (requirements for the revision of the Framework Education Program), distance learning due to the pandemic (2020-2022), an increased number of students with a different native language (e.g. Federičová, 2019; Laufková & Novotná, 2018; Münich & Protivínský, 2022; Zatloukal et al., 2021, 2022; Žlábková & Rokos, 2013), and the requirements of strategic documents such as Strategy 2020 and Strategy 2030+ (cf. Fryč et al., 2020; Ministry of Education, Youth and Sports, 2014).

The emphasis on changing the assessment of student learning and outcomes in Czech primary and lower secondary schools has sparked numerous discussions not only among curriculum developers, legislators, and strategists at the national level, but also among teacher educators, teachers in schools, parents, and researchers. The reason for this is the complexity of this change, both in terms of the demands on teachers' professional competencies for assessment and the challenges of implementing the change process among various stakeholders, including parents. As previous research findings show, parents are interested in detailed, high-quality, and understandable final assessments of student learning outcomes that include information about their children's learning process (Guskey & Bailey, 2001; Klement, 2019; Swan et al., 2014; Tomanová, 2014). Nevertheless, the traditional approach to assessment still dominates in primary and lower secondary schools in the Czech Republic. This approach has been criticized for more than a decade by the authors of the OECD report on educational assessment in the Czech Republic (Santiago et al., 2012). According to the findings of the Czech School Inspectorate, teachers still prefer the grading form (five-point numerical scale) for both ongoing and final assessment of student learning outcomes (Zatloukal et al., 2022).

In view of the strategic educational goals of the Czech Republic (Strategy 2030+), the results published in the annual reports of the Czech School Inspectorate, and the revision of the curriculum (Framework Education Program), which emphasizes a competency-based, personality-development model of education, there is a growing need to change the approach to student assessment in the Czech Republic. This change is not a minor adjustment, but rather a significant innovation of the final assessment (report card) offered to teachers, parents and their children, which would also transform the ongoing assessment.

The research is a part of the project "Development and Implementation of a Pilot Design for Report Card: Innovation in Final Assessment of Primary and Lower Secondary School Students in Connection with Comprehensive Developmental Assessment," supported by the Technology Agency of the Czech Republic (TL05000360), which attempted such a change in the concept of the report card. In the course of the three-year project, we observed the process of implementing the report card innovation in four primary and lower secondary schools in the Czech Republic. The present study is a part of this project. Our study is unique within the European context, building on the pilot American study by Swan et al. (2014) and responding to the call by Trumbull and Gerzon (2013) for explicit accounts of teachers' professional development in the area of formative assessment and its impact in local contexts. Furthermore, we identify opportunities for further research on this topic in both European and global contexts.

The data used in this paper were originally collected for a master's thesis (Grombířová, 2023). The thesis focused on the responses of all participants in the educational process, including teachers, students, and parents, to the innovative report card. For the purposes of this paper, the dataset was narrowed down to focus solely on the data gathered from teachers and school leadership. The data was then subjected to analysis with a particular emphasis on the change process involved in implementing the innovative assessment of student learning outcomes.

In this paper we follow the process of change in one selected primary and lower secondary school. The aim of the study is to describe the development of the process of changing the final assessment of student learning outcomes at this selected primary and lower secondary school and its impacts on school practice. We frame this within the concept of the Five Stages in the Decision-Making Process of Innovation (Rogers, 1983). Spurná (2019) conducted an analysis of existing theories that address the process of change. The analysis revealed that Rogers' (1983) theory provides a comprehensive and holistic view of the change process. Spurná (2019) highlights that the key strength of Rogers' theory lies in its comprehensive understanding of the innovation transfer process, which Rogers (1983) conceptualized as a form of social change occurring under specific conditions and within a defined timeframe. In the present case, this refers to the process of implementing the innovative report card in the school, which was the focus of our research.

1 Legislative requirements for student performance assessment in the Czech Republic and the need for change

Primary and lower secondary education (ISCED 1 and ISCED 2) in the Czech Republic is mostly organized within a single system of nine-year primary and lower secondary schools. These schools are divided into first and second levels, with the age of students typically ranging from 6 to 15 years (European Commission, 2024).

The mandatory requirements placed on teachers in the Czech Republic with regard to the assessment of students are set forth in legislative acts. These include Act No. 561/2004 Sb. and Decree No. 48/2005 Sb., which provide the legal framework within which teachers are obliged to operate. These requirements elucidate the extant concept of student assessment in the Czech Republic, which is reflected in the attitudes and practices of teachers. According to the findings of research and evaluations, there is a clear necessity for change. This is based on the assessment of the level of achievement of the learning outcomes for each individual subject in a school's educational program, as outlined in §14 of Decree No. 48/2005 Sb. The legislation does not provide a comprehensive framework for ongoing assessment. The assessment of behavior and academic performance in individual subjects on the report card (§§15–16, Decree No. 48/2005 Sb.) is of particular focus, with teachers permitted to utilize narrative feedback, grading, or a combination of both. This system was implemented in the Czech Republic in 2005. Subsequently, in 2012, the OECD conducted a review of the assessment and evaluation system in the Czech Republic. With regard to student assessment, several significant issues were identified, including the dominance of summative assessment, the early stages of development of assessment for learning, insufficient attention to student self-assessment, inadequate development of professional competencies in assessment during teacher pre-service training, limited communication of information to parents, and the absence of a comprehensive system for assessing learning and students learning outcomes in schools (Santiago et al., 2012). These findings indicate the need for reform in the field of student assessment.

The need for change also applies to the final assessment of student learning outcomes. The final assessment is conveyed by a document called the report card, which is issued to students at the middle and end of the school year. The content and form of the school report card are regulated by Decree No. 3/2015 Coll. Despite significant changes in both the planned and implemented curriculum for primary and lower secondary schools (standards-based curriculum), which now include key competencies as the main goal of primary and lower secondary education in the Czech Republic (Ministry of Education, Youth and Sports, 2023), the concept and form of the report card

have remained largely unchanged over the past 100 years (Bartošová & Fryč, 2011). It is therefore not surprising, as data from the Czech School Inspectorate show, that the assessment of student learning and outcomes in the Czech Republic still relies primarily on a grading system (Zatloukal et al., 2022), and that the broader implementation of ongoing formative assessment in teaching has not been successful.

Overall, it can be summarized that the legal requirements for student assessment are not comprehensively defined in a way that would serve as a tool to support student learning, contribute to their school success, and be in line with current assessment trends in the European and global context (Kratochvílová, 2013). Our assessment lags behind the general global trend, which distinguishes between outcome – what a student knows and can do at a given time, providing sufficient information about the strengths and weaknesses of the student's performance in various areas, indicating areas for improvement; process – how the student achieved these learning outcomes, including their effort, responsibility, work habits, etc., and a very important aspect of assessment – the student's progress (cf. Black & Wiliam, 1998; Earl, 2003; Guskey & Bailey, 2001; Guskey, 2011; Pollard et al., 2014; Tomanová, 2014).

Report cards in the Czech Republic do not offer a corresponding form that is aligned with ongoing student assessment and appropriately uses the information that can be conveyed through the report card to support further student development (Grombířová, 2023; Guskey, 2011). This situation presents a significant challenge – to offer teachers, parents and their children an innovation in the final assessment (report card), which would also transform the ongoing assessment.

2 JINAK report card – part of comprehensive developmental assessment of student learning and outcomes

The JINAK report card represents a significant innovation in the approach to the final assessment of student learning outcomes in primary and lower secondary schools in the Czech Republic. This initiative follows the transformation of teaching and learning culture in the Czech context, which was associated with the events of 1989 and marked the beginning of numerous changes following the political regime change. This paper does not aim to describe in detail this innovation's development, form, or content. However, to understand the subsequent text and the fundamental change it represents in the context of summative assessment on report cards, we will briefly describe the JINAK report card.

The JINAK report card combines two assessments. First, a summative assessment of student learning outcomes based on the curriculum outputs of the assessed subjects (analytical criteria-based assessment on a written scale), along with optional subject grading (holistic assessment of the subject) twice per school year (January, June) with written descriptive feedback. This feedback provides students and parents information about the student's behavior and learning outcomes in relation to the expected learning outcomes of the educational subjects in the school's educational program, highlighting the student's strengths and weaknesses, including recommendations for further development. Second, a heteronomous assessment by teachers combined with autonomous assessment by the student (Kratochvílová et al., 2024).

The report card uses a combination of descriptive feedback, criterion-based and numerical assessments. For the first time in the history of report cards in the Czech Republic, a criteria-based assessment of expected learning outcomes is officially used. This is the first time that subject outcomes are assessed analytically with the support of assessment criteria (see Appendix 1 and 2), rather than holistically with a single grade (MUNI PED, 2024).

The JINAK report card offers two completely different four-level scales for assessing behavior and expected learning outcomes, in contrast to the traditional five-level grading scale long used in the Czech Republic. The level of acquired competencies is newly assessed on a four-level scale according to established criteria and indicators. Descriptions expressing the frequency and degree of the students' independent application of the relevant skill are used to convey the achieved level for clarity and easier understanding.

The report card template uses abbreviations of words from the scale. Movement on the scale is represented by an arrow, visually emphasizing the processual aspect of learning, the student's progress, and their potential for further development (growth mindset). The assessment of character education subjects introduces another innovation: it does not include the grading scale or any other scale, but only provides space for descriptive feedback. The goal of this assessment is to support the student's development. For the first time, the report card includes a space for the student's self-assessment. The last page of the report card is dedicated to this purpose. Schools can create their own instruments for pupil self-assessment, so the format remains the choice of the school and teachers (Kratochvílová et al., 2024).

The above changes are so fundamental that the report card cannot be implemented in practice without systematic and continuous work by teachers and students in line with the report card concept. This means that the report card serves as a catalyst for changing the culture of teaching and learning. The report card provides comprehensive information on student educational progress and achievement for a given period of time on a biannual basis.

It is part of an assessment system that is used systematically and continuously over time. The report card takes a holistic view of the student, focusing on the assessment of knowledge, skills and attitudes relevant to life in relation to the key competences set out in the Framework Education Program and integrating formative and summative purposes. Its aim is not only to summarize students' learning outcomes, but also to offer opportunities for further development and to guide students towards taking responsibility for their own learning. This change in the culture of assessing student learning and outcomes is encapsulated in the concept of comprehensive developmental assessment, as defined by Kratochvílová (2012). This concept is adapted to the context of the Czech educational system and curriculum and emphasizes the following: collaboration among educational stakeholders in assessment, the use of different types and forms of assessment, the integration of formative and summative assessment, and a focus on educational goals as the objects of assessment. Comprehensive developmental assessment combines assessment for learning, assessment as learning, and culminates in assessment of learning (Earl, 2003; Hutchinson & Young, 2011; Pollard et al., 2014; Wiliam, 2011b). It follows that research monitoring the implementation of this innovation in schools is an important opportunity to understand the process of change in schools and to support its dissemination.

The implementation of the report card in schools addresses the discrepancy between reform efforts to support the development of students' personalities, the long-standing unsatisfactory state of assessment in our schools, and the traditionally applied and valid report card forms according to Decree No. 3/2015 Sb. The results of the project for developing the report card have attracted the interest of institutions, including the Ministry of Education, Youth and Sports of the Czech Republic, the National Pedagogical Institute, and the Czech School Inspectorate. These institutions have expressed interest in innovating legislation in line with the strategic and curricular goals in the Czech Republic.

3 Process of changing the final assessment of student learning outcomes

The process of changing the final assessment of student learning outcomes is a complex one. Fullan and Miles (1992) aptly characterized the relationship between change and the process of change within the educational system, stating that "education is a complex system, and changing it is even more complex." They defend this assertion, stating that even with seemingly simple changes, "the number of components and their interrelationships involved in the change is enormous and overwhelming." Furthermore, educational

changes are challenging due to the fact that a significant proportion of current practice is based on tradition rather than solid evidence of effectiveness. This is particularly evident in the case of grades and report cards (Guskey & Bailey, 2001). This is also the case in the Czech Republic (Münich & Protivínský, 2018).

The focus of experts in this field is on the process of change and its management, both in general (e.g. Daft, 1989; Donnelly et al., 1997; Kotter, 2009, 2015; Rogers, 1983; Urban 2003; Veber et al., 2016) and in the context of education (Fullan, 2007; Fullan & Miles, 1992; Goldenberg, 2004). We understand the process of change as a learning process, as aptly described by Fullan and Miles (1992, p. 745): "Change is a process of coming to terms with new personal meaning, and thus, it is a learning process." The learning process is closely linked to the professional development of teachers (Avalos, 2011; Fullan & Hargreaves, 1992) and the formation of a learning community (Shulman & Shulman, 2004; Stoll et al., 2006), which has the potential to emerge and evolve during times of change, thereby providing significant support to the successful implementation of innovations. In their study, Fullan and Hargreaves (1992) established a direct link between professional development of teachers and the process of innovation implementation, emphasizing the need for increased attention to this issue. As demonstrated by Avalos (2011), this transition is occurring gradually. Trumbull and Gerzon (2013) offer a comprehensive overview of the professional development of teachers in the field of formative assessment, emphasizing the complexity and importance of this area. It is insufficient for professional development to rely on a few workshops; it must be based on meaningful content, providing teachers with opportunities to practice, collaboratively reflect and refine their methods. In this context, a change in the final assessment of student learning outcomes can be seen as a complex process, which requires the implementation of an ongoing professional development program that is carefully designed and meaningful. This includes a process of reflection and experience exchange among the participating teachers, which serves to foster mutual enrichment and facilitate the gradual creation of a shared system that aligns with the school's overall vision and local context. The implementation of an innovative report card represents a process of change, a learning process within the school environment. All participants in the educational process – including school leadership, teachers, students, and parents – are involved in the change of the final assessment of student learning outcomes. The extent of involvement is subject to contextual factors that influence the process. It is therefore crucial to acknowledge the distinctive nature of each educational system, as well as the specific characteristics of individual schools and their current circumstances and capabilities.

It is nevertheless possible to identify certain patterns and key aspects that recur in the processes of change. The description of the stages of the change process or its key aspects is addressed in the following works: Bentley (2010), Daft (1989), Donnelly et al. (1997), Fullan and Hargreaves (1992), Kotter (2009, 2015), Lazarová (2005a, 2005b), Lunenburg (2010), Mareš (2018), Rogers (1983), Pol et al. (2013), Senge (2007), Snyder (2017), Tichá (1999), Veber et al. (2016) and Zimmerman (2006). A common element in the concept of the change process is the emphasis on the implementation of change, which Daft (1989) discusses in detail. The author presents strategies for successfully managing change, including identifying the actual need for change, the development of an idea that meets that need, the provision of support to organizational leadership, the implementation of change in a gradual manner, the planning of strategies to overcome resistance to change, the formation of teams to address different parts of the change, and the involvement of a change supporter (e.g. a volunteer who fully trusts the change, a change proposer, or a change promoter). Daft's (1989) recommendations are strongly indicative of the significance of change management and teachers' attitudes toward change. Similarly, Berkovich (2011) posits that teachers' attitudes toward change are a significant determinant of its success. The negative attitudes of Czech teachers toward change are addressed in the works of Lazarová (2005a, 2005b) and Mareš (2018). These authors, along with Průcha (2002), consistently recommend that this issue should receive more attention. The significance of attitudes towards innovations during their implementation is also reflected in Rogers' (1983) work, which emphasizes the compatibility of innovations with values and professional beliefs.

In our paper, we draw on Rogers' (1983) concept of the change process, which he refers to as the "Model of the Innovation-Decision Process." According to the author, the process of implementing innovation has five stages: (1) Knowledge; (2) Persuasion; (3) Decision; (4) Implementation; and (5) Confirmation. These stages determine the course and development of the change process. The results are then discussed in relation to these stages.

4 Methodology

4.1 Research Questions

The aim of the study is to examine how the process of implementing an innovative final assessment (report card) of student learning outcomes is developing at a selected primary and lower secondary school and what its impacts are on school practice. **The main research question is: How is the process of implementing the innovative report card developing in**

a selected primary and lower secondary school and what are its impacts on school practice?

This primary research question is complemented by three sub-research questions (SRQs) formulated to capture the development of individual concepts:

- SRQ 1: What contextual factors significantly influence the process of change regarding the final assessment of student learning outcomes?
- SRQ 2: How are teachers' attitudes toward the change in final assessment of student learning outcomes evolving, and how do they perceive the impacts of this innovation?
- SRQ 3: How do teachers describe their work with the innovative report card?

4.2 Research design

To comprehensively investigate the evolution of the change process during the implementation of the innovative report card, we employed a qualitative research design in the form of a case study. Unique cases are relatively uncommon, distinctive, and innovative. Mareš (2015) defines the case study as a descriptive and exploratory approach to research, which is used to provide a detailed and comprehensive description of a real-life phenomenon within its natural context. As defined by Creswell and Poth (2018), a case study is a research method that examines a specific case, bounded by a real-life system over time, through detailed data collection from multiple sources. The resulting data is then used to describe the case in question. The case study is an appropriate approach for mapping processes (Creswell & Poth, 2018; Mareš, 2015; Sedláček, 2014) and allows for an in-depth understanding of complex social phenomena in relation to contextual factors (Maxwell, 2013; Sedláček, 2014; Yin, 2018). In terms of case study typology, the process of implementing the innovative report card at the school under study exhibits the characteristics of a unique case (cf. Creswell & Poth, 2018; Mareš, 2015; Sedláček, 2014).

4.3 Research sample

The school in which the research was conducted was purposively selected from among schools piloting the innovative report card, based on the characteristics of the school. It is the largest comprehensive school (representing the category of large schools in the project, with more than 600 pupils), which has been implementing various innovations for a long time, but its experience with formative assessment is applied mainly at the primary school level, especially in the lower grades. The selection of this school allows us to understand the process of change in a large school where many different participants influence the process.

This is a fully organized primary and lower secondary school with a long-standing interest in educational innovation. It has bilingual classes and uses the "Step by Step" method at the primary school level. The school's requirements for the assessment and self-assessment of students in the primary level classes are described in detail and clearly in the School Code. Students in Years 1 and 2 are assessed through verbal and written feedback, while students in Years 3 to 9 are assessed using a traditional grading scale. The rules for the assessment and self-assessment of students at the secondary level are also described in detail in the School Code and are consistent across all grades, reflecting a traditional approach to assessment. The aim of the school leadership in taking part in this research was not only to support the change in assessment in the first and second grades, but also to extend this change to the higher grades. The current format of the report card in the first and second grades (narrative feedback in the form of letters) no longer fully satisfies the teachers. The school expects the following outcomes from its participation in the research: a) to extend the assessment reform to all grades and to unify the assessment rules within the school, b) to change the collection of evidence of learning for ongoing assessment and its use in the report card, c) to make the report card clearer for parents, d) to change the report card system across the school.

The school leadership discussed participation in the pilot project with the teachers as a group and individually. Participation was not mandatory, and the development of the report card was time consuming for the teachers; for example, teachers had to write two report cards over three semesters: the official one, valid according to the current legislation, and the experimental one. For these reasons, the school leadership decided to include selected primary and lower secondary grades in the pilot of the innovative school report card, which also became the research sample.

Participants in the study included primary school teachers ($n = 3$), lower secondary school teachers ($n = 20$) teaching in classes piloting the innovative report card, and school leadership staff ($n = 2$). Their previous training in formative assessment consisted of short online webinars.

The characteristics of the teachers who were interviewed individually, including their training in formative assessment, are presented in Appendix 3. In Appendix 4 the list of other teacher pseudonyms we used is presented.

4.4 Data collection

The data collection process was conducted in three stages over a period of 14 months, from January 2022 to March 2023. This was done in conjunction with the development of the innovative report card form, as depicted in Figure 1.

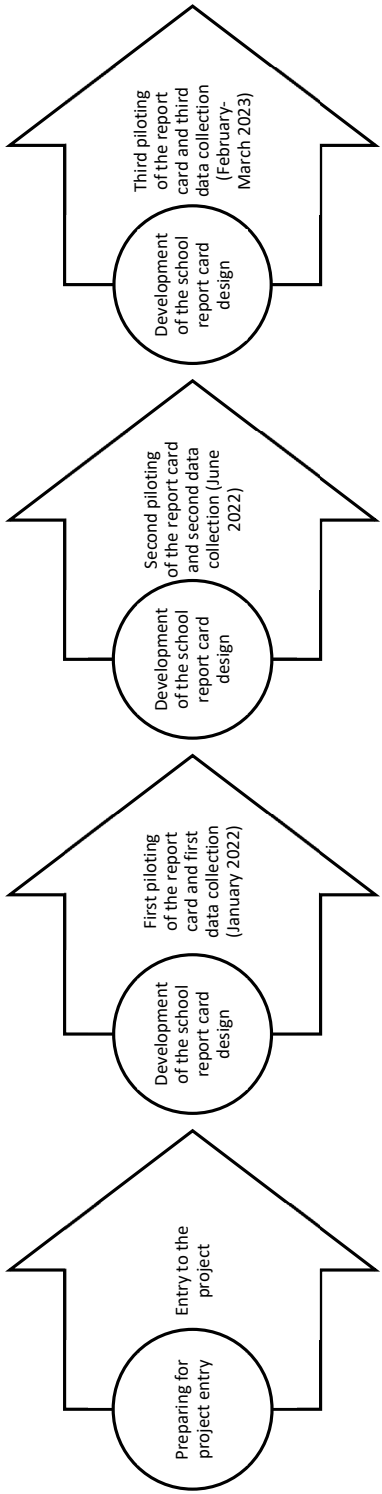


Figure 1
Stages of data collection

The initial two data collections were conducted via semi-structured group interviews. In each instance, the interviews were conducted by two members of the project team. The interviews included teachers from both the primary and secondary school levels, as well as school leadership personnel ($n = 25$). The objective of the interviews was to gain insight into the teachers' experiences with the initial and subsequent iterations of the innovative report card, and to elucidate the process of implementing the change. The interviews were conducted in person and recorded for subsequent analysis.

In the third stage of data collection, individual semi-structured in-depth interviews were conducted with two primary school teachers, two lower secondary school teachers, and the principal. The objective was to facilitate a profound reflection on the teachers' experiences with the third iteration of the innovative report card and to delineate the evolution of the implementation process. The participants were selected by the principal at the request of the research team. The interviews were conducted according to the same interview scheme, via the online platform ZOOM, and were recorded.

Data from the group and individual interviews became the data corpus and provide evidence of the progression of each stage of the change process. A total of 532 minutes of recordings were obtained. All interview recordings were transcribed in accordance with established transcription rules (Švaříček, 2014). All data were anonymized, and respondents were informed about the purpose, content, and anonymization of the interview prior to its commencement. The respondents were afforded the opportunity to inquire about the process and content of the interview. In the case of individual interviews, respondents were also assured of anonymity through email correspondence prior to the interviews. Written consents were included in the project documentation.

4.5 Data analysis

All the data were analyzed in several stages after the completion of data collection. During the analysis we carefully distinguished between the different stages of the change process. First, we proceeded inductively. In the first phase, we coded all interviews using the method of open coding – pencil and paper – and applied in vivo codes (Šeďová, 2014). As Merriam (2009) asserts, at this stage the researcher remains open to all possible explanations, which allows for a comprehensive representation of the change process and captures all phenomena that occur within it. In total, we generated 595 codes. In the second phase, we grouped the inductively generated codes into subcategories ($n = 31$) and categories ($n = 25$) based on thematic relationships. Then, in the third phase, we grouped the categories into main categories ($n = 3$), which addressed the sub-research questions leading to the main

research question (Šeďová, 2014). Finally, we organized the main categories, categories and subcategories into schemas, which helped us to interpret the data. Through this interpretation, we answered the research questions and captured the development of the change process from the perspective of primary and secondary school teachers and school leadership. The results indicated that the change process in the school reflected Rogers' (1983) model of the stages of the change process; therefore, in discussing the results, we proceeded deductively, relating our findings to his Model of the Innovation-Decision process. We also discussed our findings with regard to the findings of other authors.

4.6 Study limits

It is important to note that our study has a few limitations. The first is that in the third phase of data collection, the school principal selected the respondents. She was asked to choose teachers with positive and negative attitudes. However, the final choice was up to her, so we don't have the views of all the teachers. The second limitation is that this study focuses only on the perspectives of teachers and school leadership. Although data reflecting the views of other actors in the educational process—such as parents and students —were part of the wider research, we focused here more closely on the process of change among the main stakeholders implementing the innovative report card, using data from group and individual interviews. There are other ways we were able to collect more data, such as conducting classroom observations, analyzing the completed innovative report card forms, or getting quantitative data from questionnaires.

5 Research findings

The research questions are answered through the interpretation of the data. The development of the change process was observed from the perspective of primary and lower secondary school teachers.

5.1 Development of the context of the change process

Given the unique nature of each change process occurring in a school and the distinct characteristics of each school, our initial focus is on describing the contextual factors that enter into and influence the change process.

During the study, primary and lower secondary school teachers delineated a number of contextual factors that emerged during the three stages of the change process. These factors were grouped into the categories "Factors Facilitating the Change Process" and "Factors Hindering the Change Process."

5.1.1 Factors Facilitating the Change Process

The data analysis indicates that the entire process of implementing the innovative report card was significantly supported by: a) A system of teaching organization: knowledge of students; b) Experience with formative assessment: long-term work with formative assessment and student self-assessment in previous years; c) Collegial sharing: from chance meetings to regularity; and d) management of the change implementation process.

a) System of teaching organization: Knowledge of students

Teachers identified knowledge of students as a subcategory that significantly influences the acceptance of innovation in practice. Primary school teachers consistently highlighted the significance of the teaching organization at the primary school, which determines the more extensive time spent with students. An extensive time frame permits teachers to gain a comprehensive understanding of their students, enabling them to conduct both formal and informal diagnostic assessments, which are crucial for a thorough grasp of the students' abilities and needs. The results of these diagnostic activities provide substantial support for formative assessment and for gathering information to formulate the final assessment in the innovative report card.

> *Without this, I would have been unable to accomplish the task. From September to January, I simply recorded the names of the students, their competencies, and their participation in group work, using a system of tick marks. I also made notes regarding my interactions with the students. This notebook serves as a repository for observations, which are then graded. This was the most important aspect for me. (Jane)*

The organization of teaching in lower secondary school differs significantly from that in primary school. Teachers in lower secondary school are typically less familiar with of their students and lack the data necessary for providing descriptive feedback on innovative report cards. Teachers Susan and Anne indicated that their role as class teachers facilitated their ability to work with the innovative report card in comparison to other teachers who do not have as frequent contact with their students.

> *It was beneficial to have a deep understanding of the students. (Susan)*

The results underscore the disparate pedagogical structures at the primary and lower secondary levels, which shape the extent of teacher-student contact and influence the receptivity to change.

b) Experience with formative assessment: Long-term work with formative and self-assessment of students in previous years

Previous experience was a significant factor for primary school teachers in working with the innovative report card. The teachers placed a high value on their extensive experience with the use of both formative and summative

written feedback in their pedagogical practice. Previously, report cards were written in the form of letters, representing narrative feedback. According to the teachers, these were often lengthy and lacked structure. The transition to a different innovative report card was not as challenging for them.

Indeed, the process of filling in the report card was relatively straightforward. (Mary)

The introduction of the innovative report card prompted the educators to streamline their narrative feedback, reduce its length, and engage in more profound reflection on how to capture the essence of their students' achievements more succinctly, as evidenced by their responses in the interviews.

Additionally, both Mary and Jane had considerable experience with a similar four-point scale utilized on the innovative report card. A comparable scale was employed for both student assessment and self-assessment in an assessment tool, namely a book designed for the systematic documentation of assessments and self-assessments of students.

That made the process somewhat more straightforward, as we had previously utilized similar rating scales. Two of us worked with that particular student book. (Jane)

Teachers at the lower secondary level had limited experience with formative assessment or any other assessment scale beyond the official grading scale. However, the school leadership provided ongoing professional training for teachers to acquire new knowledge and skills to support formative assessment and provide quality feedback in the classes. This professional development was positively received by the teachers:

It helps me. The seminars on formative assessment are helpful... (Susan)

The results clearly show that previous experience or lack thereof with formative assessment in practice, narrative and descriptive feedback, and using scales other than the standard grading scale proved to be significant factors that either supported or hindered teachers' work with the innovative report card.

c) Collegial sharing: From chance meetings to regularity

A prominent theme in the change process for primary school teachers was the importance of collegial support through the sharing of ideas and practice. Primary school teachers initiated meetings to discuss challenging topics, including behavior assessment and systematic evidence collection.

Gradually, they began to share their experiences on a regular basis. In the third stage of the project, collegial support and the associated development of guidelines and materials for student assessment and self-assessment at the school became systematic and regular.

For me, this was one of those things where we had to sit down, we had to put it together, and we talked about how it should work. So, it was more like professional guidance… (Mary)

In the third stage of the change process, the school leadership stated that they also supported collegial sharing at the lower secondary level, but were aware of the difficulty of finding free time for collegial meetings at the secondary level. An expert committee was created to share experiences with the innovation. Lower secondary school teachers were inspired by the guidelines for the primary level, as mentioned by the principal:

> *At the lower secondary school, we created a committee, an expert committee, and teachers meet there and create a guideline for assessment for the lower secondary school as well.*

During the implementation of the innovative report card in practice, it became clear how important it is to share experiences and support each other when difficulties arise. Collegial sharing evolved into regular meetings and the gradual development of an internal assessment system.

d) Management of the change implementation process

The implementation of the innovative report card involved a significant degree of change management within the school. This was carried out by the school leadership, comprising the principal and her deputies. Initially, the school leadership did not disseminate the methodological support received from the project team to the teachers. However, over time, teachers came to recognize and appreciate the role of the management in creating conditions conducive to the implementation of the innovative report card. These conditions included: a) Support for collegial sharing: Regular teachers' meetings**;** b) Involvement of a respected staff member: A member of the teaching staff was engaged to provide support for change management among teachers. c) Ongoing professional development in formative assessment: Professional development activities were conducted for lower secondary school teachers. d) Sufficient time allocation: Adequate time was allotted for the explanation of the innovative report card to students, which facilitated their understanding of the new report card form. e) Communication with parents: Information about the innovative report card was disseminated to parents through various channels, including instructions for communication, sample emails, sharing of videos from the innovative report card website, and posting information on the school website.

5.1.2 Factors hindering the change process

A data analysis revealed that the process of implementing the innovative report card was hindered by several factors: a) Novelty factor: lack of advance knowledge of the final form of the innovative report card and the assessment scale; b) Teaching organization system: lack of familiarity with the students; c) Assessment of key competencies (behavior area); d) Instructional guideline for the innovative report card as a support tool**.**

a) Novelty factor: Lack of advance knowledge of the final form of the innovative report card and assessment scale

The initial version of the innovative report card was developed gradually between June and December 2021. The form for the innovative report card was distributed to schools in early December, and teachers began filling in the first innovative report cards in January. This timing significantly impacted the implementation of the innovation.

The date, according to the teachers, resulted in a lack of understanding of the content of the innovative report card form and prevented the integration of assessed objectives into monitoring student development during the first half of the 2021/2022 school year and planning the teaching process. As one teacher stated:

> *I had a problem with getting it so late. I spent half the summer imagining what I would do and what evidence of learning I would have. (Mary).*

Teachers were thus confronted with the reality (despite having been kept informed of the development of the template) that its design, particularly the assessment criteria, were a new element for teachers for which they had not purposefully gathered evidence of learning. In the second and third stages of the change implementation, primary school teachers accepted the form. Lower secondary school teachers, however, did not fully accept the assessment criteria, as they differed from their original expectations.

Another significant innovation in the report card form was the assessment scale. While some primary school teachers were already familiar with the scale, others found it challenging. For instance, Lily, a primary school teacher, initially perceived it as a novel element with which she had no prior experience. However, she gradually became more comfortable with it, stating:

> *I don't really work with the scale: not yet mastered, partially acquired, fully acquired. It doesn't quite fit for me personally.*

In later stages of the innovative report card implementation, primary school teachers no longer focused on the scales and did not mention them in the interviews, suggesting that they had accepted the new system. Lower secondary teachers did not mention the scale in the interviews.

b) Teaching organization system: Lack of familiarity with the students

While primary school teachers perceived knowing their students as an advantage, the process of changing the final assessment of student learning outcomes at the lower secondary school level was repeatedly hindered by a lack of knowledge about their students. This issue was further compounded by the limited time allotted for some subjects, which ranged from one to two lessons per week. In some cases, lower secondary school teachers were unable to recall the names of all their students in multiple classes.

I am a physical education teacher; I do not recall the names of my students. (Petr)

Consequently, lower secondary school teachers lacked sufficient high-quality evidence of student learning and adequate materials to formulate the final assessments for the innovative report card.

Moreover, I lacked the evidence of students' learning. (Betty)

In contrast, class teachers had a different experience, as previously mentioned. They concurred that an understanding of their students was beneficial when filling in the innovative report card.

c) Assessment of key competencies (behavior area)

The conflict regarding the assessment of key competencies (referred to as "Behavior" on the innovative report card) was particularly prominent among teachers at the lower secondary school level. For these teachers, the introduction of key competencies on the innovative report card represented a significant change from previous practice. Prior to this, teachers at the lower secondary school had not employed this option, which presented a conflict and potentially a threat. Despite the Czech Republic's primary and lower secondary education curriculum (Framework Education Program) emphasizing the acquisition of key competencies (Ministry of Education Youth and Sports, 2023), the system lacks the capacity to support their systematic assessment. Moreover, their assessment is not mandatory on the graded report cards. Teachers may utilize an additional form for narrative feedback, wherein they describe students' levels of key competencies. However, this practice is exceedingly uncommon, and the school had no prior experience with this form of competency assessment. Teacher Martin posited that assessment reform should be accompanied by curriculum reform. He strongly opposed the assessment of key competencies:

No one from the implementation team in 2005 ever mentioned that key competencies should be quantified, ranked, or measured in any way. It seems completely absurd to me.

Consequently, the team largely rejected the assessment of key competencies. This rejection was also due to some teachers' lack of clarity regarding the place and significance of key competencies within the curriculum hierarchy.

Therefore, these key competencies should be cross-curricular. How is a teacher to recognize that? I must confess that I am uncertain. (Petr)

They repeatedly rejected the possibility of a meeting among several teachers to discuss the assessment of key competencies:

In view of the number of other teachers involved, a discussion would be necessary in order to have a full and accurate assessment. However, neither I nor the other teachers have the time to devote to such a discussion. (Mary)

In addition, the time required to discuss each student in every class would be considerable. This is an issue that primary school teachers did not address. Could their responses be influenced by the opinions of their colleagues from lower secondary school?

d) Instructional guideline for the innovative report card as a support tool

In order to facilitate the implementation of the innovative report card, the project team developed instructional guidelines for teachers. The instructional guidelines were disseminated to the school leadership for distribution among teachers at the beginning of December, in advance of the preparation of the first term innovative report cards (January 2022). The instructional guidelines provide comprehensive guidance for the completion of the innovative report card, including descriptions of assessment scales and illustrative examples. However, due to an organizational oversight, teachers did not receive the instructional guidelines from the school administration during the initial stage of the change process, which became a significant limiting factor, as openly described in interviews following the second stage of the pilot:

We didn't really have it in the first semester… (Jane).

In the second stage of the change process, teachers were given access to the instructional guidelines and began to utilize them. Primary school teachers mainly used the methodology for the assessment of behavior, a new element on the innovative report card, as previously mentioned.

I primarily utilized the instructional guidelines for those competencies. (Lily)

Additionally, some teachers at the lower secondary school level began using the instructional guidelines. For instance, Megan closely adhered to the instructional guidelines, which validated her approach.

I was pleased to find that my practice was similar to that described in your instructional guide, which gave me confidence that I had followed the correct approach.

Additionally, Susan provided commentary on the instructional guidelines in the final stage. According to her, the instructional guidelines should include a detailed description of the process of creating an assessment system and its implementation in the school, rather than merely providing support for filling out the final assessment form of the innovative report card. Despite the sharing of experiences among teachers from various schools during the implementation of the innovative report card, Susan offered a critical observation:

This is not a critique of this specific innovative report card or this pilot, but rather a comment on the lack of guidance provided by the university. As a large school with at least 14 teachers involved in each class, we require assistance in implementing this in practice.

We summarize the factors facilitating and hindering the process of change in the final assessment of student learning outcomes in Table 1, providing an overview of these factors and their positive (+), negative (–), early (0) or absent (X) charge in relation to whether they were primary or lower secondary school teachers at each stage of the change process.

Table 1
An overview of the factors influencing the change process and their dynamics

Factor	Primary school teachers (Stage 1)	Primary school teachers (Stage 2)	Primary school teachers (Stage 3)	Lower secondary school teachers (Stage 1)	Lower secondary school teachers (Stage 2)	Lower secondary school teachers (Stage 3)
Instructional Guidelines for the Innovative Report Card as a Support Tool	0	+	+	–	–	–
Assessment of Key Competencies (behavior area)	0	+	+	–	–	–
Teaching Organization System: Lack of Understanding the Students	X	X	X	–	–	–
Lack of Advance Knowledge of the Final Template of the Innovative Report Card and Assessment Scale	–	X	X	–	X	X
Management of the Change Implementation Process	0	+	+	0	+	+
Collegial sharing	0	+	+	0	0	+
Experience with Formative Assessment	+	+	+	0	+	+
Understanding of the students	+	+	+	X	X	X

In summary, the development of contextual factors in the change process was related to the duration of the change process. The development of their influences was particularly evident in the first and second stages of the change process, and it continued to intensify in the third stage.

5.2 The development of teachers' attitudes towards the change in the final assessment of student learning outcomes and the perceived impacts of this change

According to Berkovich (2011), teachers' attitudes toward change are crucial in the process of its implementation. During the interviews, primary and lower secondary teachers described their attitudes, which fell into two categories: a) Attitudes Toward Student Assessment; b) Attitudes Toward the Innovative Report Card. These attitudes, together with their overall conception of the innovative report card and its implementation in school practice, had a wide range of positive impacts and contributed significantly to fostering a productive culture of teaching and learning. In the following section, we present the development of these attitudes and the perceived impact of the change, focusing on: a) The development of attitudes and perceived impacts of the change among primary school teachers; b) The development of attitudes and perceived impacts of the change among lower secondary school teachers.

a) The development of attitudes and perceived impacts of the change among primary school teachers

From the beginning, the primary teachers agreed on the need for change in the final assessment of student learning outcomes. In the first stage of the change process, they expressed positive attitudes towards the innovative report card, recognizing the benefits of the innovation and predicting that long-term use of the report card would lead to pupils understanding its content and the assessments it contained.

> *I think if a child starts using it from the first year and knows what each section means, it will be good, but it also includes the work before and after... (Lily)*

In the second stage, primary teachers' attitudes remained consistently positive and became more entrenched. They embraced the comprehensive concept of the innovative report card, were willing to collect evidence of learning to formulate final assessments, and indicated that long-term use of the report card would improve its quality. The innovation was gradually integrated into their work system. The only exception was Mary, who disagreed with the idea that the report card should include recommendations for the student's future development:

> *I don't think that belongs in the final report card. I just want to tell them where they are now and I can tell them about their development separately or sometime during the process.*

However, recommendations for further development are required by Decree No. 48/2005 Coll. as part of the narrative feedback.

In the third stage of the change process, the attitudes of primary school teachers remained stable and positive, which was reflected, for example, in their appreciation of the structure of the innovative report card. Despite this long-standing positive attitude, teachers in the third stage were concerned

about the use of the innovative report card in grades 3-5, where several teachers are involved in a class. This led to challenges in sharing information about how students performed, particularly regarding behavior:

> *It felt like the problems they have in lower secondary school, like getting assessments from colleagues who teach in my class. Forcing them to do what I'm doing. Or somehow getting it out of them. (Mary)*

Primary school teachers have accepted the complex concept of the innovative report card and their attention has gradually shifted to student self-assessment and the wider positive impacts of working with the innovative report card. These impacts, which can be summarized as impacts on the primary school, impacts on the lessons, and impacts on individual teachers, are shown in Table 2.

Table 2
Summary of the impacts of the implementation of the innovative report card in primary school

Impacts on primary school	**Impacts on the lessons**	**Impacts on teachers**
extension of collegial support	support for self-regulated learning	better consistency in the collection of evidence of learning
vision and development of materials for graded student self-assessment	improving work with pupil self-assessment	keeping written records of formative assessment and pupil self-assessment
changing the frequency of assessment (School Code)	supporting work with objectives in teaching	a comprehensive view of the pupil
vision to create a system of assessment at primary school	focusing teaching on the development of pupils' key competencies	systematization of assessment work
collaborative development of assessment materials and methodologies	promoting formative assessment	individualization of written feedback
regular assessment meetings (once a month)		better consistency in the collection of evidence of learning
extension of collegial support		keeping written records of formative assessment and pupil self-assessment
		improving the quality of written feedback on report cards
		acceptance of assessment scales

b) The development of attitudes and perceived impacts of the change among lower secondary school teachers

Attitudes towards the need for change in the final assessment of student learning outcomes varied among lower secondary teachers. While some teachers agreed with the need for change, others felt that the traditional grading system was sufficient, as described by Petr:

For me the numbers are enough. That's all I need.

However, in terms of their attitudes towards the innovative report card, lower secondary teachers shared similar views. Their main concern was the time it would take to fill in the final assessment into the innovative report card form and to assess key competencies. These negative attitudes were influenced by several factors: some lower secondary teachers did not agree with keeping detailed evidence about students, assumed that parents would not be interested in the innovative report card, and did not accept the standardized criteria-based assessment derived from the national curriculum, as they had their own ideas about assessment criteria. As Karin explained:

We all agreed – I didn't assess any of the students in this class because I don't teach there, I teach in the other class, but it doesn't matter. We agreed with the Czech teachers that the categories of grammar, literature and composition don't suit us at all. ... We initially proposed some more specific categories and we think that the ones we have formulated, although we could still discuss them, better reflect what we actually do in class... (Karin)

In the second stage of the change process, teachers in lower secondary schools had a strongly negative attitude towards the innovative report card. They disagreed with the behavior assessment, felt that the innovative report card was only suitable for small schools, and objected to the time and cognitive demands of formulating the final assessment in the innovative report card form:

Yes, so that it's not more formal than the actual space for teaching. (Megan)

This negative attitude towards the innovative report card was explained by lower secondary teachers with various arguments – they mentioned time demands, school size and overall workload. Betty was an involuntary participant in the innovative report card pilot:

... just at the personal request of someone I respect.

Petr, on the other hand, felt that publishing the educational objectives on the school's website was sufficient and that there was no need to communicate them further to students and parents.

That's what the school plans on the website are for, the child knows, right. Those who want to know will find it, and those who don't care won't be interested.

Lower secondary school teachers have not yet embraced the innovative report card. However, in the third stage of the change process, Susan and Anne characterized its positive impacts. The implementation process supported professional development, work with formative assessment, the systematization of student learning outcome assessments, and the development of descriptive feedback.

> *...so, I really tried to make sure that some of the feedback was more descriptive and not judgmental. (Anne)*

In addition, the innovative report card provided them with a comprehensive view of the student.

The attitude of primary and lower secondary teachers towards the innovative report card was stable throughout the pilot phase. Primary school teachers, together with the school leadership, adopted a positive long-term attitude toward the innovation. Based on this attitude, primary school teachers have successfully integrated the innovative school report into their work system and have gradually started to diffuse it to other primary school teachers through collaborative teacher groups focused on student assessment. As they described, this integration has had extensive positive impacts. Lower secondary teachers, on the other hand, maintained a long-term negative attitude towards the innovation, which has not yet been overcome, resulting in the innovation being largely rejected by lower secondary teachers. Nevertheless, the experience with the process of change of the final assessment of student learning outcomes has had positive impacts.

5.3 Development of work with the innovative report card from the perspective of teachers

The work with the innovative report card is specific and requires a high level of professional competence of the teachers. This section captures the development of work with the innovative report card, based on descriptions from teachers at different stages of the change process: a) Development of work with the innovative report card from the perspective of primary school teachers; b) Development of work with the innovative report card from the perspective of lower secondary school teachers.

a) Development of work with the innovative report card from the perspective of primary school teachers

In the first stage, primary and lower secondary teachers completed the innovative report card in its entirety, except for the student self-assessments. A wide and diverse range of evidence of learning proved essential for working with the innovative report card. In this stage, primary school teachers used diagnostic records, student self-assessments, narrative and descriptive

feedback, and intuition, especially for assessing key competencies. The new element of the innovative report card – the expanded behavioral assessment (key competencies) and the absence of instructional guidelines may have led primary school teachers to rely primarily on intuition to assess key competencies:

We didn't get a chart for these competencies. So, I kind of deduced it intuitively. (Mary)

In the second stage, the primary school teachers expanded their portfolios of ways to collect evidence of learning. There was a change in the source of behavioral assessment; Mary no longer cited intuition, but memory:

I think about it and then I say: "yes, Honzík..." (They also worked on understanding the content and assessments filled in by students and their parents in the innovative report card.)

In the third stage, the primary school teachers used the complete form of the innovative report, including the student self-assessment and the possibility of parental feedback. They continued to systematize the collection of evidence of learning, focusing on working with students to help them understand the innovative report card. As a result, the primary school teachers integrated the rating scale into the continuous assessment and self-assessment of the students, gradually worked on assessment tools, mapped the students' attitudes towards the innovative report card, and explained their personal position to the students.

b) Development of work with the innovative report card from the perspective of lower secondary school teachers

Teachers at the lower secondary school lacked a sufficient amount of evidence of learning. A new element for them in the innovative report card was the space for written descriptive feedback, which complemented the criterion-based assessment. This innovation was mostly rejected by lower secondary teachers. As some of them mentioned, they copied texts for several students and justified it with the number of students in the class:

Well, everyone got the same sentence. Everyone got the same sentence because I didn't have time to do it differently. (Martin)

Teachers at the lower secondary school continued to work with the innovative report in a similar way during the second stage of the change process. Some began to use criteria developed by subject committees in the school to collect evidence of learning, while others used grades. Many reported that they did not have enough evidence of learning:

And I didn't have the materials. (Betty)

Lower secondary school teachers also found it increasingly challenging to formulate descriptive feedback for the innovative report card.

In the third stage of the change process, Susan focused on collecting evidence of learning, which consisted primarily of teacher assessment records, student self-assessments, and peer assessments. Susan also worked with student progress:

> *...because I approached it responsibly and really wrote personalized comments in this report card... Each child has something different from me; I don't copy anything. I compare it to the first semester...*

However, Susan was an exception at the lower secondary level, as most lower secondary teachers did not systematize their collection of evidence of learning during the three stages of the pilot.

For both primary and lower secondary teachers, having sufficient quality learning evidence was crucial to working with the innovative report card in all three stages of the change process. Primary teachers worked systematically on collecting learning evidence and gradually started to use the full innovative report card form. In contrast, lower secondary teachers often filled out the innovative report card formally, paying little attention to its quality and potential.

5.4 Summary

In the present case study, we focused on a two-year change process – the implementation of the innovative report card concept into school practice. We identified key contextual factors that significantly influenced the implementation process, either positively (understanding of students, experience with formative assessment, collegial sharing, change management) or negatively (timing of receiving the template, lack of understanding of students, assessment of key competencies (behavior), insufficient work with the instructional guidelines). We then focused on teachers' attitudes towards the innovation, which played a crucial role in its acceptance (primary school) or non-acceptance (lower secondary school). These attitudes were influenced by teachers' perceptions of the final form of the innovative school report card and its complexity. We then characterized the development of work with the innovative report card and found that having sufficient high-quality evidence of learning was essential for its use. Changing the final assessment of student learning outcomes takes time, but as shown, it can have a significant positive impact on the productive culture of teaching and learning. In the case of the primary school, we have identified impacts that demonstrate this. Furthermore, even though the innovative report has not yet been adopted in the lower secondary school, it has had a positive impact on this culture.

6 Discussion and conclusion

The study characterized the development of the process of change in the final assessment of student learning outcomes and its perceived impacts on a selected primary and lower secondary school from the perspectives of teachers and school leadership. The school was chosen because of its characteristics as one of the pilot schools for the innovative report card. The focus was on the contextual factors of the change process, the attitudes of the teachers, the work with the innovative report card and the impacts of the innovative report card. The change of the report card is a new phenomenon in the Czech Republic. Current professional and public discussions related to the strategic intent of Strategy 2030+ indicate that this is a highly relevant issue. In this chapter, we summarize and discuss our research findings in relation to the model of the Innovation Decision-Making Process (Rogers, 1983), comparing them with findings from other researchers.

Rogers (1983) states that in the first stage of the change process, **Knowledge** is essential. Teachers seek and acquire information about the innovation, its advantages and disadvantages: What is it? How does it work? Why does it work? This step preceded the initiation of the implementation process of the innovative report card into school practice. The point of conflict became the assessment of key competences and the requirement for descriptive feedback, specifically for lower secondary school teachers. The findings revealed a discrepancy between the state curriculum's requirements (key competencies as one of its objective domains), legislative requirements (the absence of legislative foundation for assessing key competencies), and the innovative demand to monitor the development of key competencies and officially assess them on the report card. Lower secondary school teachers did not consider the assessment of key competencies to be a priority educational objective. In contrast, primary school teachers demonstrated greater acceptance of the assessment of key competencies. The crucial role of pedagogical content knowledge in transforming assessment practices at the primary school level is also a primary finding of Jones and Moreland (2005) case study. The importance of knowledge in implementing innovation in school environments was similarly highlighted by Avidov-Ungar and Eshet-Alkakay (2011) and Roehrig and Kruse (2005). In their study, Avidov-Ungar and Eshet-Alkakay (2011) examined the implementation of innovative technologies in primary schools and concluded that the coexistence of a learning organizational culture within the school, along with teachers' high level of technological-pedagogical knowledge, plays a crucial role in fostering positive attitudes and successful implementation. Roehrig and Kruse (2005) investigated the importance of beliefs and knowledge in adapting reform-based curriculum in high school chemistry class. Their findings align with

ours in showing that content knowledge was a critical factor for the successful adaptation of innovation. Additionally, knowledge of change as a process is essential, as demonstrated in management-focused works (e.g., Daft, 1989; Donnelly et al., 1997; Kotter 2009, 2015; Rogers, 1983; Urban 2003; Veber et al. 2016). Findings by Washington and Hacker (2005) show that managers who understand the change process are more open to changes and are likely to be invested in successful implementation. Analysis of our data showed that the school leadership made efforts to support the change process. The support system was gradually developed during the process of implementing the innovation.

In the second stage, which Rogers (1983) calls **Persuasion**, the process involves exploring how the innovation works, with teachers experimenting or imagining how it might work: What will happen if I accept it? They seek support, reassurance, and answers from others and various sources, and they evaluate the pros and cons of the innovation. This stage coincides with the first stage of introducing the innovative report card into school practice. After the initial stage of getting acquainted with the innovation, primary and lower secondary teachers were engaged in identifying their needs and ideas about the innovative report card. Kotter (2009, 2015) considers this stage of the change process (creating a sense of urgency for change) to be the most important. While primary school teachers began to consider refining the current assessment system, lower secondary school teachers, even in this first stage, began to oppose the innovation for various reasons – the assessment of key competencies, the rejection of the proposed assessment criteria, the time-consuming process of gathering evidence and completing the innovative report card form. This phase was significantly affected by the lack of information summarized in the methodology, which was supposed to support them but was not provided in time. It becomes clear that the different phases of the change process are interrelated and that the information phase cannot be underestimated.

The third stage of the change process, according to Rogers (1983), is **Decision**. The individual engages in activities that lead to acceptance or rejection of the innovation (active rejection: considering acceptance, experimenting; passive rejection: outright rejection of the innovation). The attitude of the teachers towards the innovation is important for this decision. In the school, this stage occurred during the second experience with the innovative report card. For primary school teachers, their positive attitude was the main issue. For lower secondary school teachers, their negative attitude towards the innovative report card significantly deepened. The research results clearly indicate that both the successful and unsuccessful implementation, adaptation and diffusion of the change were primarily influenced by these positive/negative attitudes of teachers towards the

innovation, reflecting their value orientation. Lazarová (2005a) in her review also points to the crucial importance of value orientation to the innovation during its introduction. The process of changing teachers' beliefs and attitudes in educational change is operationalized by Guskey (1985) in A New Model of Teacher Change, in which he presents that sustainable change in teaching practices occurs only after teachers' beliefs and attitudes have changed. Their change occurs as a result of improvements in student learning outcomes that result from changes in teaching practices. In the context of our findings, it is possible that lower secondary school teachers have not yet seen these changes in student learning outcomes and therefore their attitudes have not changed over the time period we investigated.

The review showed that researchers investigate teachers' beliefs in the context of change (Ham & Dekkers, 2019; Haney et al., 2002; Lebak, 2015; Richards et al., 2001; Roehrig & Kruse, 2005) rather than their attitudes (Anghelache & Benţea, 2012; Maskit, 2011). Anghelache and Benţea (2012) examined high school teachers' level of resistance to change and Maskit (2011) focused on primary school, junior high school and high school teachers' attitudes towards pedagogical change at different stages of their professional development. They unanimously concluded that resistance to change is related to the stage of a teacher's career – teachers at later stages of their career tend to have more negative attitudes towards change. Our findings differ significantly from the conclusions of these authors. In the research we present, these were lower secondary school teachers, regardless of their career stage. This confirms the words of Mareš (2018) and Průcha (2002) that we lack more comprehensive research on this area, with only sub-research studies available, which is a great opportunity for further educational research.

The fourth stage, according to Rogers (1983), is **Implementation**. The innovation is put into practice and answers to further questions are actively sought. Needs such as sharing, mentoring, and stability arise. This stage began at the primary school during the third stage when the innovation was successfully adapted and disseminated throughout the primary school. This successful adaptation was supported by the management of the change process, including strategies for effectively managing it – gradually implementing the change, creating teams to work on different parts of the change, and involving a change supporter, as recommended by Daft (1989) in his theory. This approach resulted in significant positive impacts. These positive impacts are considered one of the key findings of our research, as they highlight the specific benefits of introducing the innovation – the innovative report card – into school practice. Gradually, the primary school began to move into the fifth stage, which according to Rogers (1983) is **Confirmation**. In this stage, there is stabilization, but there can also be regression due to conflicting reports about the innovation. In the primary

school, doubts arose based on the experiences of colleagues in the lower secondary school and the problem of not sharing information among colleagues. They considered whether it would be possible to use the innovative report card in the upper classes. In the lower secondary school, the fourth and fifth stages of the change process did not occur as a result of negative attitudes, although positive impacts of the innovation on school practice were identified.

The context of the research suggests that changing the (final) assessment is a current challenge for the Czech education system, even in the context of the ongoing so-called Big Revision of the Framework Educational Programs by the National Pedagogical Institute (2024), which should include the elimination of grading in grades 1-3 of primary school. The innovative report is presented as a suitable alternative in this case. But the problem goes deeper. Despite the curricular reforms, there have been no significant changes in school assessment (Straková et al., 2013; Zatloukal et al., 2020, 2021). This change in the assessment of learning and student learning outcomes, which will support the learning process and the quality of each student's life, will be a long and challenging process, but it can bring significant positive impacts, as the presented study shows. The negative effects of the current assessment system in the Czech Republic have been described by Federičová (2019) and Münich and Protivínský (2018, 2022). Therefore, there is evidence for the need to change the system of assessment learning and student learning outcomes in the Czech educational context.

The innovated report card has already influenced the change to Decree No. 48/2005, which explicitly added the criterial assessment. An update to Decree No. 3/2015 is currently being prepared to include the innovative report card. The legislative requirements for the assessment and the possibility of using the innovative report card in practice are therefore gradually changing. The experience described in this paper can be an inspiration for teachers and schools in the implementation of the innovative report card in practice, which is the result of the transformation of continuous assessment that supports the learning of our students.

The success of the change requires:

a) Rethinking not only student assessment but also the organization of instruction, especially in the lower secondary schools. Block scheduling could provide more time for monitoring student progress, supporting self-assessment, and collecting evidence. In addition, it is worth considering the use of teachers who could teach multiple subjects in a single class, thereby gaining a more comprehensive understanding of students across subjects.

b) Thoroughly understanding the form of the innovative report card, and understanding both the criteria and the rating scales. In-depth study of the instructional guideline is essential. Knowledge and understanding are the foundation for implementing change in practice. The initial stage of learning about the innovation should not be underestimated.

c) Carefully planning the management of the change process, creating time for collegial sharing, which should be planned and systematic. School leadership must demonstrate the need for change and provide management for others.

d) Identifying the benefits of the change and its impact as a source of justification for implementation, alongside the weaknesses perceived by teachers, such as the lack of time to collect evidence of learning, particularly in subjects with low weekly time allocation.

e) Reflecting continuously on the experience of the change and responding promptly to situations that affect the still unstable attitudes of teachers to the detriment of the change. Responding to these situations and strengthening the impacts of the change in the short and long term are essential.

Acknowledgments

In this paper, partial results from a case study conducted as part of the thesis are presented: Grombířová (2023). Teachers', students', and parents' reactions to the change in final assessment – case study [Master's thesis, Masaryk University]. Archive of Theses, MUNI. https://is.muni.cz/auth/th/od2o2

The research was supported by the project Technology Agency of the Czech Republic Development and Implementation of a Pilot Design of Report Card: Innovation of Final Assessment of Primary and Lower Secondary School Students in Connection with Comprehensive Developmental Assessment (TL05000360).

References

Anghelache, V., & Benţea, C. C. (2012). Educational changes and teachers' attitude towards change. *Procedia-Social and Behavioral Sciences*, *33*, 593–597. https://doi.org/10.1016/j.sbspro.2012.01.190

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10–20. https://doi.org/10.1016/j.tate.2010.08.007

Avidov-Ungar, O., & Eshet-Alkakay, Y. (2011). Teachers in a world of change: Teachers' knowledge and attitudes towards the implementation of innovative technologies in schools. *Interdisciplinary Journal of E-Learning and Learning Objects*, *7*, 291–303. https://doi.org/10.28945/1525

Bartošová, J., & Fryč, J. (2011). *Vysvědčení jako součást dějin školství: od 18. století po současnost*. National Pedagogical Museum and Library of J. A. Comenius.

Bentley, T. (2010). Innovation and diffusion as a theory of change. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 29–46). Springer.

Berkovich, I. (2011). No we won't! Teachers' resistance to educational reform. *Journal of Educational Administration*, *49*(5), 563–578. https://doi.org/10.1108/09578231111159548

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. NFER-Nelson.

Creswell, J. W., & Poth, Ch. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches*. SAGE Publications, Inc.

Daft, R. L. (1989). *Organization theory and design*. West Publishing Company.

Decree No. 3/2015 Coll., Decree on Specific Educational Documents (2015). https://www.aspi.cz/products/lawText/1/83287/1/2

Decree No. 48/2005 Coll., Decree on Primary Education and Certain Requirements for Fulfilling Compulsory School Attendance (2005). https://www.aspi.cz/products/lawText/1/59591/1/2

Donnelly, J. H. jr., Geibson. J., & Ivancevich. J. M. (1997). *Management*. Grada.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

European Commission. (2024, 7th October). *Eurydice – Czechia*. Eurydice. https://eurydice.eacea.ec.europa.eu/national-education-systems/czechia/overview

Federičová, M. (2019). *Dopady známek na vysvědčení na životní rozhodnutí žáků*. Economics Institute of the Czech Academy of Sciences.

Fryč, J., Matušková, Z., Katzová, P., Kovář, K., Beran, J., Valachová, I., Seifert, L., Běťáková, M., Hrdlička, F., & et al. (2020). *Strategie vzdělávací politiky České republiky do roku 2030+*. Ministry of Education, Youth and Sports.

Fullan, M. G. (2007). *The new meaning of educational change*. Routledge.

Fullan, M., & Hargreaves, A. (1992). Teacher development and educational change. In M. Fullan & A. Hargreaves (Eds.), *Teacher development and educational change* (pp.1–9). Routledge.

Fullan, M. G., & Miles, M. B. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan*, *73*(10), 745–752.

Goldenberg, C. N. (2004). *Successful school change: Creating settings to improve teaching and learning*. Teachers College Press.

Grombířová, J. (2023). *Reakce učitelů, žáků a rodičů na změnu závěrečného hodnocení – případová studie*. [Diplomová práce, Masarykova univerzita]. Archiv závěrečných prací MUNI. https://is.muni.cz/th/od2o2

Guskey, T. R. (1985). Staff development and teacher change. *Educational leadership*, *42*(7), 57–60.

Guskey, T. R. (2011). Five obstacles to grading form. *Effective Grading Practices*, *69*(3), 16–21.

Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Corwin Press.

Ham, M., & Dekkers, J. (2019). What role do teachers' beliefs play in the implementation of educational reform?: Nepali teachers' voice. *Teaching and Teacher Education*, *86*, 102917. https://doi.org/10.1016/j.tate.2019.102917

Haney, J. J., Lumpe, A. T., Czerniak, C. M., & Egan, V. (2002). From beliefs to actions: The beliefs and actions of teachers implementing change. *Journal of Science Teacher Education*, *13*(3), 171–187. https://doi.org/10.1023/A:1016565016116

Hutchinson, C., & Young, M. (2011). Assessment for learning in the accountability era: Empirical evidence from Scotland. *Studies in Educational Evaluation*, *37*(1), 62–70. https://doi.org/10.1016/j.stueduc.2011.03.007

Jones, A., & Moreland, J. (2005). The importance of pedagogical content knowledge in assessment for learning practices: A case-study of a whole-school approach. *Curriculum Journal*, *16*(2), 193–206. https://doi.org/10.1080/09585170500136044

Klement, L. (2019). *Evropské školy jako inspirativní model vzdělávání evropského občana v 21. století se zaměřením na první stupeň ZŠ* [Disertační práce]. Centrum profesních studií.

Kotter, J. P. (2009). *Vědomí naléhavosti: první a nejdůležitější krok realizace změny*. Management Press.

Kotter, J. P. (2015). *Vedení procesu změny*. Management Press.

Kratochvílová, J. (2011). *Systém hodnocení a sebehodnocení žáků, zkušenosti z České republiky i Evropských škol*. MSD.

Kratochvílová, J. (2012). Aktivní spoluúčast žáka při hodnocení – zdroj inspirace rozvoje osobnosti žáka a pokládání základů zodpovědnosti za kvalitu svého života. In H. Lukášová (Ed.), *Proměny pojetí vzdělávání a školního hodnocení* (s. 151–180). Asociace waldorfských škol ČR.

Kratochvílová, J. (2013). *Inkluzivní vzdělávání v české primární škole: teorie, praxe, výzkum*. Masaryk University.

Kratochvílová, J. (2015a). Pupil assessment within the context of inclusive education in primary schools. In *International multidisciplinary scientific conference on social sciences snd arts SGEM 2015* (pp. 773–780). STEF92 Technology.

Kratochvílová, J. (2015b). Pupil self-assessment and assessment of the teacher–two closely related processes. *Przegląd Pedagogiczny*, *1*, 19–29. https://doi.org/10.34767/PP.2015.01.02

Kratochvílová, J., Spurná, M., & Bařina Barochová, E. (2024). *Souhrnná výzkumná zpráva*. Masaryk University. https://vysvedcenijinak.ped.muni.cz/media/3618480/tacr-odborna-zprava-fin-2512024.pdf

Laufková, V., & Novotná, K. (2018). Školní hodnocení z pohledu žáků. *Orbis scholae*, *8*(1), 111–127.
https://doi.org/10.14712/23363177.2015.8

Law No. 561/2004 Coll., the Education Act, regulates preschool, primary, secondary, higher vocational, and other types of education in the Czech Republic (2004).
https://www.aspi.cz/products/lawText/1/58471/1/2

Lazarová, B. (2005a). Osobnost učitele a rezistence vůči změně. *Studia paedagogica*, *53*(10), 109–122.

Lazarová, B. (2005b). Psychologické aspekty profesionálního rozvoje učitelů: rezistence vůči změně. *Pedagogika*, *55*(2), 102–118.

Lebak, K. (2015). Unpacking the complex relationship between beliefs, practice, and change related to inquiry-based instruction of one science teacher. *Journal of Science Teacher Education*, *26*(8), 695–713.
https://doi.org/10.1007/s10972-015-9445-0

Lunenburg, F. C. (2010). Forces for and resistance to organizational change. *National Forum of Educational Administration and Supervision Journal, 27*(4), 1–10.

Mareš, J. (2015). Tvorba případových studií pro výzkumné účely. *Pedagogika*, *65*(2), 113–142.

Mareš, J. (2018). Odpor učitelů vůči změnám ve škole a ve školství. *Pedagogika*, *68*(2), 173–200.
https://doi.org/10.14712/23362189.2018.1090

Maskit, D. (2011). Teachers' attitudes toward pedagogical changes during various stages of professional development. *Teaching and Teacher Education*, *27*(5), 851–860.
https://doi.org/10.1016/j.tate.2011.01.009

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach.* Sage Publications.

Merriam, S. B. (2009). *Qualitative research: A guide to design and imlementation*. Jossey-Bass.

Ministry of Education, Youth and Sports. (2014). *Strategie vzdělávací politiky České republiky do roku 2020.* Ministry of Education, Youth and Sports.

Ministry of Education, Youth and Sports. (2023). *Rámcový vzdělávací program pro základní vzdělávání.* Ministry of Education, Youth and Sports.

MUNI PED. (2024). *Vysvědčení jinak: ukázka šablony finální verze.* Masaryk Univesity.
https://vysvedcenijinak.ped.muni.cz/ukazka-vysvedceni

Münich, D., & Protivínský, T. (2018). *Co skrývají známky na vysvědčení?* Economics Institute of the Czech Academy of Sciences.

Münich, D., & Protivínský, T. (2022). *Rozdíly v přísnosti známkování žáků a dopady na vzdělanostní aspirace.* Národohospodářský ústav AV ČR, v.v.i.

Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students.* Harlow.

National Pedagogical Institute. (2024). *Revize RVP základního vzdělávání.* National Pedagogical Institute.
https://velke-revize-zv.rvp.cz/

Pol, M., Hloušková, L., Lazarová, B., Novotný, P., & Sedláček, M. (2013). *Když se školy učí. Masaryk University.*.

Pollard, A., Black-Hawkins, K., Cliff-Hodges, G., Dudley, P., & James, M. (2014). *Reflectiveteaching in schools: Evidence-informed professional practice*. Bloomsbury Publishing.

Průcha, J. (2002). *Učitel: současné poznatky o profesi*. Portál.

Richards, J. C., Gallo, P. B., & Renandya, W. A. (2001). Exploring teachers' beliefs and the processes of change. *PAC Journal*, *1*(1), 41–58.

Roehrig, G. H., & Kruse, R. A. (2005). The role of teachers' beliefs and knowledge in the adoption of a reformbased curriculum. *School Science and Mathematics*, *105*(8), 412–422. https://doi.org/10.1111/j.1949-8594.2005.tb18061.x

Rogers, M. E. (1983). *Diffusion of innovations*. Macmillan Publishing.

Santiago, P., Donaldson, G., & Herman, J. (2012). *OECD reviews of evaluation and assessment in education*. OECD. https://www.oecd-ilibrary.org/docserver/9789264116788-en.

STUDIA PAEDAGOGICA

Studia paedagogica is a peer reviewed journal published by the Faculty of Arts at Masaryk University. The editorial board is located at the Department of Educational Sciences. The publication frequency is three issues per year.

Studia paedagogica publishes papers on education, upbringing and learning from all spheres of social life. The papers are theoretical, but mainly empirical. The journal is interdisciplinary – it covers current topics in educational research while at the same time providing scope for studies grounded in other social sciences.

The journal is indexed in international databases (CEJSH, DOAJ, EBSCO, ERA, ERIH PLUS, PKP, ProQuest, SCOPUS).

Publisher

Masaryk University, Czech Republic
This issue is published in December 2024.

Journal website

www.studiapaedagogica.cz

Contact

Department of Educational Sciences
Faculty of Arts, Masaryk University
Arna Nováka 1, 602 00 Brno
Czech Republic
e-mail: studiapaedagogica@phil.muni.cz

The journal is a continuation of SPFFBU, U series.
ISSN 2336-4521 (online)

STUDIA PAEDAGOGICA
VOLUME 29 / NUMBER 3 / YEAR 2024

Contents

Department of Educational Sciences
Faculty of Arts, Masaryk University
www.studiapaedagogica.cz
ISSN 2336-4521 (online)

MUNI PRESS

MUNI ARTS