Learning and Instruction 105 (2026) 102431

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Longitudinal and differential effects of differentiated instruction on academic achievement and subjective well-being in inclusive secondary education

Carmen L.A. Zurbriggen [a,*], Marcela Pozas [b], Karin Gehrer [c], Lena Nusser [c]

[a] *Faculty of Education, University of Fribourg, Rue St. Pierre-Canisius 21, Fribourg, 1700, Switzerland*
[b] *Faculty of Humanities, Education and Social Sciences, University of Luxembourg, Porte des Sciences 11, Esch-sur-Alzette, 4366, Luxembourg*
[c] *Leibniz Institute for Educational Trajectories, Bamberg, Wilhelmsplatz 3, Bamberg, 96047, Germany*

A R T I C L E I N F O

Keywords:
Differentiated instruction
Academic achievement
Subjective well-being
Inclusive education
Competencies

A B S T R A C T

Background: Differentiated instruction (DI) is a fundamental component of teaching quality, particularly in diverse and inclusive classrooms. While DI has been shown to enhance students' academic achievement, research at the secondary education level remains limited, and its impact on students' subjective well-being—widely recognized as essential for learning—has been largely unexplored. Moreover, effects may differ for students with special educational needs in learning (SEN-L).
Aim: This study investigated cross-sectional and longitudinal effects of classroom-level DI practices on students' academic achievement and subjective well-being in lower secondary education, incorporating both student and teacher perspectives. Additionally, it examined whether the effects of DI differ between students with and without SEN-L.
Sample: Data were drawn from a longitudinal study with three measurement occasions (Grades 6, 7, and 9). The sample comprised 4654 sixth-grade students, along with 417 German and 398 mathematics teachers.
Methods: Multilevel regression analyses with latent variables were conducted to investigate cross-sectional effects. To examine longitudinal effects, latent neighbor change models were applied.
Results: Cross-sectionally, DI showed some small positive associations with academic achievement and, particularly student-reported DI, small positive associations with subjective well-being. DI predicted some changes in both outcomes over time, albeit inconsistently and with small-to-negligible effect sizes. The findings indicate only two small differential effects for students with SEN-L compared to other students.
Conclusions: The study extends research on DI by focusing on both students' academic achievement and subjective well-being in lower secondary education, while calling for continued research into its longitudinal effects and underlying mechanisms.

1. Introduction

"Today's classrooms are typified by academic diversity" (Tomlinson et al., 2003, p. 119). This statement introducing the landmark review article on differentiated instruction (DI) remains as salient today as it was two decades ago. Initially designed to tailor teaching to specific student groups, DI steadily evolved into a major teaching approach grounded in the principle of educational equity, taking into account the diverse cognitive and learning needs of all students within a classroom (Langelaan et al., 2024; Valiandes, 2015).

DI is widely recognized as a fundamental component of effective instruction (Parsons et al., 2018). Empirical evidence, predominantly from primary education, indicates mixed but generally positive results (Graham et al., 2021). The meta-analysis by Deunk et al. (2018) revealed a small overall positive effect on students' academic achievement, particularly when DI was embedded in a broader educational framework or school reform, such as inclusive education, which explicitly aims to address learner diversity through supportive learning environments (Tomlinson, 2017).

The mixed pattern of findings may indicate differential effects or "aptitude-treatment interactions" (Cronbach & Snow, 1981; Tetzlaff

* Corresponding author. University of Fribourg, Rue St. Pierre-Canisius 21, 1700, Fribourg, Switzerland.
E-mail addresses: carmen.zurbriggen@unifr.ch (C.L.A. Zurbriggen), marcela.pozas@uni.lu (M. Pozas), karin.gehrer@lifbi.de (K. Gehrer), lena.nusser@lifbi.de (L. Nusser).

https://doi.org/10.1016/j.learninstruc.2026.102431
Received 24 October 2025; Received in revised form 7 June 2026; Accepted 2 July 2026

et al., 2025). Accordingly, several studies have suggested that effects of DI differ for low-versus high-achieving students (Karst et al., 2022) and for students with versus without special educational needs (SEN) (Nusser & Gehrer, 2020). Low-achieving students in particular appear to benefit from DI in reading (Förster et al., 2018) and mathematics (van der Scheer & Visscher, 2018). This may be due to teachers implementing DI less often for high-achieving students (Prast et al., 2015; Ziernwald et al., 2022). Moreover, Ziernwald et al.'s (2022) systematic review also suggests that both teachers and students perceive DI as supportive and motivationally beneficial. This aligns with the conclusions of Graham et al. (2021) and Smale-Jacobse et al. (2019). Both emphasize the importance of considering non-academic outcomes—such as student engagement, affect, and subjective well-being—due to their significant relevance to learning and academic development.

Compared to academic achievement, research on DI including subjective well-being or other socio-emotional characteristics remains limited, especially in secondary education (Pozas et al., 2021; Smale-Jacobse et al., 2019). This is particularly noteworthy, given that students' subjective well-being is widely recognized as both highly relevant to learning and a favorable educational outcome (Bücker et al., 2018; Hascher, 2012). At the same time, accurate teacher judgments of students' subjective well-being are fundamental for DI (Zurbriggen et al., 2025). Accordingly, Pozas et al. (2021) and Alnahdi et al. (2022) found that DI practices were positively associated with students' subjective well-being.

In both studies, teaching practices were assessed from the students' perspective—a common approach in research on differentiation (Praetorius et al., 2018). In contrast, incorporating teachers' perspective can offer additional valuable insights, especially in inclusive education (Lindner et al., 2023). As Obrovská et al. (2025) recently highlighted, studies on DI that integrate both perspectives remain notably scarce. Moreover, there is a general paucity of research examining DI within whole-classroom settings (Tetzlaff et al., 2022). Most of the evidence on the effectiveness of DI for academic outcomes comes from cross-sectional or intervention studies. This limits the ability to draw conclusions about long-term effects and may compromise ecological validity.

Against this backdrop, the present study aims to extend research on DI by examining whether DI affects students' academic achievement and subjective well-being in the context of inclusive classrooms in lower secondary education, considering both teacher and student perspectives. In addition to cross-sectional associations, longitudinal effects and potential differential effects for students with versus without SEN in learning (SEN-L) will be investigated.

2. Differentiated instruction and its effects on student outcomes

The concept of DI is referred to by various terms in the literature, including adaptive teaching and personalized learning, which are sometimes used interchangeably (Dumont & Ready, 2023) despite referring to conceptually distinct instructional approaches (Parsons et al., 2018; Tetzlaff et al., 2022). In the present study, DI is conceptualized as the planned adaptation of instructions to address student heterogeneity within a shared-classroom context, following Tomlinson's (2014, 2017) framework. Accordingly, DI refers to instructional practices addressed to the classroom group at large or to flexible subgroups of students, rather than to individual learners. This distinguishes DI from individualized instruction, which focuses on tailoring instruction to each student individually (Tetzlaff et al., 2022), and from personalized learning, a concept predominantly discussed in educational technology research (Dumont & Ready, 2023). In this context, DI combines careful progress monitoring with instructional adjustments based on students' needs, incorporating a range of instructional strategies (Pozas et al., 2019; Roy et al., 2013), such as tiered assignments, flexible grouping, peer tutoring, open educational approaches, and staggered learning materials.

Within educational research, DI is widely recognized as a fundamental component of teaching quality (Praetorius et al., 2018). By adjusting instructional content, processes, products, and learning environments, DI aims to enhance student engagement, participation, and achievement (Tomlinson, 2014). Despite strong theoretical foundations, empirical research suggests that the implementation of DI often diverges from its conceptual ideal (Langelaan et al., 2024). Several studies have documented a tendency among teachers to prioritize process-related or quantitative variants of differentiation (e.g., providing extra time or additional tasks), while learning goals often remain uniform for all students (Jager et al., 2025).

Differences in the implementation of DI across school subjects have also been observed. For example, DI practices are reported to be more commonly applied in mathematics than in language arts (Pozas et al., 2019). In contrast, Saloviita (2018) found no significant school-subject-based variation in DI implementation within Finnish comprehensive schools. DI practices also vary across school tracks and school levels. According to longitudinal analyses based on a representative sample from Germany, the use of DI practices declined during secondary education (Gehrer & Nusser, 2020).

2.1. Effects on students' academic achievement

DI is commonly expected to be linked to improved academic achievement, particularly in mathematics and reading (Parsons et al., 2018). Meta-analytic findings support this assumption while also painting a more nuanced picture. Hattie (2009) reported only modest average effect sizes for DI practices. Similarly, Seidel and Shavelson (2007) found very small positive average effect sizes for DI on cognitive outcomes ($d = 0.03$). Deunk et al. (2018) identified a small overall positive effect of DI on primary students' performance in language and mathematics, although some studies reported negligible or even negative effects on reading performance.

Focusing on secondary education, a systematic review by Smale-Jacobse et al. (2019) reported moderate positive effects ($d = 0.51$–0.74) of DI on students' academic achievement. In a longitudinal study, however, DI was found to affect students' language achievement only in the early years of lower secondary education (Gehrer & Nusser, 2020). According to a meta-analysis by Steenbergen-Hu et al. (2016), the effect of DI was nonsignificant when it was operationalized solely through grouping strategies. Overall, these findings underscore the importance of examining DI at different stages of lower secondary education and longitudinally as a comprehensive approach rather than an isolated strategy.

2.2. Effects on students' subjective well-being

The effects of DI on students' subjective well-being have rarely been examined, although school-related well-being has gained momentum in educational research—particularly due to its relevance for learning processes and academic functioning (Bücker et al., 2018). The established framework by Hascher (2012) conceptualizes subjective well-being at school as a multidimensional construct encompassing emotional and social aspects as well as students' academic self-concept.

Two of the few existing studies on effects of DI on students' well-being are in line with Hascher's framework, focusing on three aspects of subjective well-being that are particularly relevant in inclusive education (Zurbriggen et al., 2019). Both studies report generally positive findings: Pozas et al. (2021) found small positive associations with emotional well-being in school ($\beta = 0.26$), social inclusion ($\beta = 0.14$) and academic self-concept ($\beta = 0.18$), and Alnahdi et al. (2022) observed more pronounced associations ($\beta = 0.38$–0.43).

Among these three components, academic self-concept has been investigated more frequently than the others, although it is often not explicitly identified as part of students' subjective well-being. For example, Kulakow (2020) found that students in competence-based DI settings reported higher academic self-concepts than those taught through teacher-directed whole-class instruction. This finding is consistent with the meta-analysis by O'Mara et al. (2006) which demonstrated that individual counseling and feedback have positive effects on students' academic self-concept.

2.3. Teacher and student perspectives of DI

A key methodological factor that may partly explain inconsistencies in findings on DI effectiveness is the perspective from which it is assessed. Differentiation and adaptive support are evaluated either from the observer, student, or teacher perspective (Praetorius et al., 2018). Expert classroom observations are often resource-intensive and time-consuming. Students' ratings can capture the suitability of and cognitive activation elicited by instructional practices, although they have been criticized for their limited ability to distinguish among different aspects of DI (Kunter & Baumert, 2006). Conversely, teachers may have the most comprehensive understanding of their own practices, but their responses can be biased by factors such as social desirability (Tetzlaff et al., 2022). Thus, incorporating both student and teacher perspectives is recommended, as each perspective provides unique yet complementary insights into DI (Kunter & Baumert, 2006; Obrovská et al., 2025).

2.4. Differential effects of DI

Differential effects may challenge one of the core justifications for implementing DI: its potential to promote educational equity (Valiandes, 2015), at least if this conceptualization is based on the idea of "equality of outcomes" but not of "adequacy" (Dumont & Ready, 2023). Several studies have suggested that DI may yield differential effects for students with different ability levels (Förster et al., 2018; Karst et al., 2022; Roy et al., 2015). One plausible explanation is that teachers primarily use DI activities to support struggling, low-achieving students (Datnow & Hubbard, 2015) and apply it less often to high-achieving students (Prast et al., 2015), which may lead to greater learning gains for the former group. Differential effects of DI may also reflect an "aptitude-treatment interaction" (Cronbach & Snow, 1981; Tetzlaff et al., 2025). By definition, an interaction between students' ability (i.e., "aptitude") and instructional approach (i.e., "treatment") is explicitly intended when responding to students' individual differences through differentiating or adapting instruction (Dumont & Ready, 2023).

Evidence on differential effects on students' academic achievement is inconsistent, though several studies suggest stronger effects for students with low achievement levels. Förster et al. (2018) found that students with lower reading skills benefited more from DI than their higher-skilled peers, while Nusser and Gehrer (2020) reported no differential effects of tiered assignments on language achievement for students with SEN-L or low achievement, similarly to the findings of Karst et al. (2022). In mathematics, van der Scheer and Visscher (2018) reported no general effect of DI but positive outcomes for low-achieving students. These mixed findings may partly stem from methodological differences, but may also reflect variation in how DI is defined and operationalized across studies, resulting in conceptually different practices being labelled as "differentiation" (Graham et al., 2021). However, findings are contradictory for more homogenous definitions as well. For within-class ability grouping, Deunk et al.'s (2018) meta-analysis reported small negative effects for low-ability students and no significant effects for high-ability students in primary education, whereas the second-order meta-analysis by Steenbergen-Hu et al. (2016) found no differential effects in either primary or secondary education.

Regarding subjective well-being, empirical evidence of possible differential effects of DI for students with SEN-L also remains inconclusive. Kulakow (2020) showed that competence-based DI settings mostly benefitted students with low academic self-concepts, which are typically observed among students with SEN-L. This result contradicts the findings of Pozas et al. (2021), who reported that SEN status was negatively associated with academic self-concept in the context of DI, while no significant associations were identified for emotional well-being and social inclusion. Conversely, Alnahdi et al. (2022) found positive associations between SEN and DI for all three aspects of subjective well-being.

One possible explanation for such differential effects might be that teachers are more attentive to the subjective well-being of students with SEN when adapting their instructional practices, as emotional well-being, social inclusion and academic self-concept are considered important outcomes in inclusive education, particularly for students with SEN (Zurbriggen et al., 2019), which could lead to greater benefits for these students. On the other hand, teachers' professional skills may be challenged when implementing DI in the context of SEN, which could potentially affect students' subjective well-being (Goldan et al., 2022).

3. The present study

The present study aims to contribute to the existing body of research on DI by investigating cross-sectional and potential longitudinal effects on students' academic achievement and subjective well-being within real classroom contexts at the lower secondary level. We focused on inclusive secondary schools because their high student diversity amplifies the necessity of DI. Furthermore, integrating DI within a broader framework has been identified as essential for its effectiveness. In order to gain complementary insights into instructional practices, we considered the perspectives of both students and teachers.

We addressed the following research questions: Is DI associated with students' academic achievement and subjective well-being in lower secondary education (RQ1)? Does DI have longitudinal effects on students' academic achievement and subjective well-being during lower secondary education (RQ2)? Do the effects of DI differ between students with SEN-L and other students (RQ3)? Based on previous findings, we anticipated that DI would have positive effects on both outcomes, especially in earlier years of lower secondary education. Drawing on theoretical assumptions and empirical evidence, we hypothesized that DI would have positive longitudinal effects on both outcomes and more pronounced effects for students with SEN-L than for other students. Nevertheless, the direction of possible differential effects remains an open empirical question.

4. Method

4.1. Sample and procedure

This study draws on data from the longitudinal project *Inclusive Education in and after Lower Secondary Schools in Germany* (INSIDE; Schmitt et al., 2023). The INSIDE study invited inclusive secondary schools to participate that reported having at least one student with SEN enrolled. Inclusive schools in Germany are characterized by a student body with diverse learning profiles, while the structural conditions of and conceptual approaches to inclusive education remain heterogeneous.

A total of 4657 sixth-grade students (48.7% female) from 766 classes across 246 schools provided consent to participate and were followed throughout lower secondary school. The study was carried out with two successive cohorts that began in 2019 and 2020, respectively. Data collection took place at three time points: in Grades 6 (T1), 7 (T2), and 9 (T3), each scheduled at the end of the school year. Due to financial reasons, no measurements were conducted in Grade 8.

The survey took place during regular lessons and was administered on-site by trained personnel. Students were informed about the procedure and given instructions for completing the competence tests and the questionnaire. Subject teachers for German and mathematics were also invited to participate and asked to complete a questionnaire on their teaching practices.

At T1, the students' mean age was 12.6 years ($SD = 0.6$). In total, 11.8% reported that German—the majority language—was not their primary language, and 8% had been formally diagnosed with SEN in the learning and/or language domains. Approximately 35.0% of the sample came from an academic family background, defined as having at least one parent with a master's degree.

4.2. Measures

4.2.1. DI from student perspective

At T1 and T3, students reported separately on the teaching practices of their German and mathematics teachers using four items for each subject (e.g., "Our German/math teacher focuses on each student individually"; Gebhardt et al., 2014). Responses were given on a scale from 1 ("do not agree at all") to 4 ("totally agree"). Students' reports reflect perceived DI within the classroom, capturing observable aspects of Tomlinson's (2014, 2017) framework, including process-related aspects, learning-environment adaptations, and aspects of student autonomy and interest (for more information: Supplement, Tables A1 & A2). The internal consistency of the scale was good for both measurement occasions and both domains (German: $\omega = 0.77$ | 0.81; mathematics $\omega = 0.80$ | 0.80). Longitudinal measurement invariance across T1 and T3 was established (Table B1). ICC ranged from 0.13 to 0.22, indicating a substantial amount of variance that could be attributed to the classroom level (Table B5).

4.2.2. DI from teacher perspective

Additionally, at T1 and T3, German and mathematics teachers reported on their teaching practices regarding DI with five items each (e.g., "I give students tasks of varying difficulty depending on their level of competence"; Theisel, 2014). Responses were given on a Likert-scale ranging from 1 ("never") to 4 ("very often"). In line with Tomlinson's (2014, 2017) framework, we focused on process-related aspects of DI, learning-environment adaptations, and aspects of student autonomy and interest. The selected items corresponded conceptually to the student-reported DI items but assessed teachers' implemented practices from an instructional perspective (Tables A3 and A4). The internal consistency of the scale was acceptable for both measurement occasions and both domains (German: $\omega = 0.66$ | 0.79; mathematics $\omega = 0.71$ | 0.72). Measurement invariance was established across T1 and T3 (Table B2). Overall, there was relatively high stability of teaching staff, with many teachers remaining with the same class (German: 69.7%; mathematics: 61.5%).

4.2.3. Student academic achievement

Academic achievement was assessed through reading comprehension and mathematical competence tests administered at each measurement occasion. These standardized tests emphasized functional application in real-life contexts, capturing not only domain-specific knowledge but also cognitive processes, such as reasoning, inference, and problem-solving. Using an anchor item design, competence scores were linked over time to model developmental trends and growth trajectories (Weinert et al., 2019). Responses were scaled using item response theory, resulting in weighted likelihood estimates (WLE; Warm, 1989). Because of the wide range of competence levels within inclusive classes, several booklets with varying item difficulties were administered.

Reading comprehension. The reading comprehension test included four or five different text types serving distinct communicative functions. The corresponding tasks required various cognitive processes, such as locating information, drawing inferences, and reflecting textual content (Gehrer et al., 2013). Reliabilities were good, with EAP/PV reliability ranging from 0.80 for Grade 6 to 0.79 for Grade 7 and 0.75 for Grade 9. ICCs ranged from 0.23 to 0.25 (Table B5).

Mathematical competence. The standardized test assessing mathematical competencies was based on everyday problem scenarios across multiple content domains, including Quantity, Change and Relationships, Space and Shape, and Data and Chance (Neumann et al., 2013). The 20-item test also measured cognitive components related to mathematical reasoning and operations. Reliabilities were good, with EAP/PV reliability ranging from 0.77 for Grade 6 to 0.82 for Grade 7 and 0.80 for Grade 9. ICCs ranged from 0.24 to 0.32 (Table B5).

4.2.4. Subjective well-being

Students' subjective well-being at school was assessed at each measurement occasion using the German version of the *Perceptions of Inclusion Questionnaire* (PIQ; Venetz et al., 2015), which captures three sub-dimensions: emotional well-being (i.e., school-related emotional well-being), social inclusion (i.e., positive social relationships with classmates), and academic self-concept (Zurbriggen et al., 2019). Each subdimension was measured with three items rated on a 4-point Likert scale ranging from 1 ("not at all true") to 4 ("certainly true"). A three-factor model showed excellent model fit across all three measurement occasions. Internal consistency values ranged from $\omega = 0.76$ to $\omega = 0.89$, and ICCs ranged from 0.04 to 0.12 (Table B5). Longitudinal measurement invariance of the subjective well-being construct was established (Table B3).

4.2.5. Control variables

Additionally, relevant individual characteristics associated with students' outcomes as well as those variables associated with non-participation in Grade 7 and Grade 9 (Table B7) were included as control variables.

Gender. Teachers reported students' gender (0 = male, 1 = female). Missing information was replaced with student or parent reports.

SEN-L. Teachers indicated whether students had an official SEN status in the area of learning and/or language (0 = no, 1 = yes).

Primary language. Teachers reported the primary language spoken by students at home (0 = German, 1 = other language).

Educational background. Parental education information was coded using the CASMIN scale (Brauns et al., 2003).

Basic cognitive skills. Students' basic cognitive abilities were assessed at T1 using the deductive reasoning subtest of the *Berlin Test of Fluid and Crystallized Intelligence* (BEFKI; Wilhelm et al., 2014). The number of correctly solved items was used as a proxy for basic cognitive skills ($M = 8.23$, $SD = 3.34$).

Specific learning difficulties. Teachers indicated whether students had a diagnosis of specific learning difficulties (e.g., dyscalculia; 0 = no, 1 = yes).

Grades. Schools reported students' grades in the subjects German and mathematics (1 = highest grade through 6 = lowest grade).

School type. Sampling information includes the school type of the respective schools. We differentiate between one-track schools, multi-track schools, and gymnasia.

4.3. Analyses

Analyses were conducted using M*plus* 7.4 (Muthén & Muthén, 2017). Syntax is provided in the supplement. Missing data were handled with the full-information maximum likelihood (FIML) approach.

To address RQ1, we conducted multilevel structural equation modelling to account for the hierarchical data structure (students nested within classes). We adopted a stepwise modelling strategy, beginning with separate Level-1 (student) and Level-2 (classroom) models, followed by a two-level model that simultaneously incorporated predictors from both levels. To examine whether associations differed across distinct stages of lower secondary education, we estimated separate multilevel models for Grade 6 and Grade 9 for all outcomes. In all models, teacher and student reports of DI served as key independent variables. We also included students' gender, SEN-L status, primary language, educational background, specific learning disabilities, basic cognitive skills and school type as covariates to ensure that observed

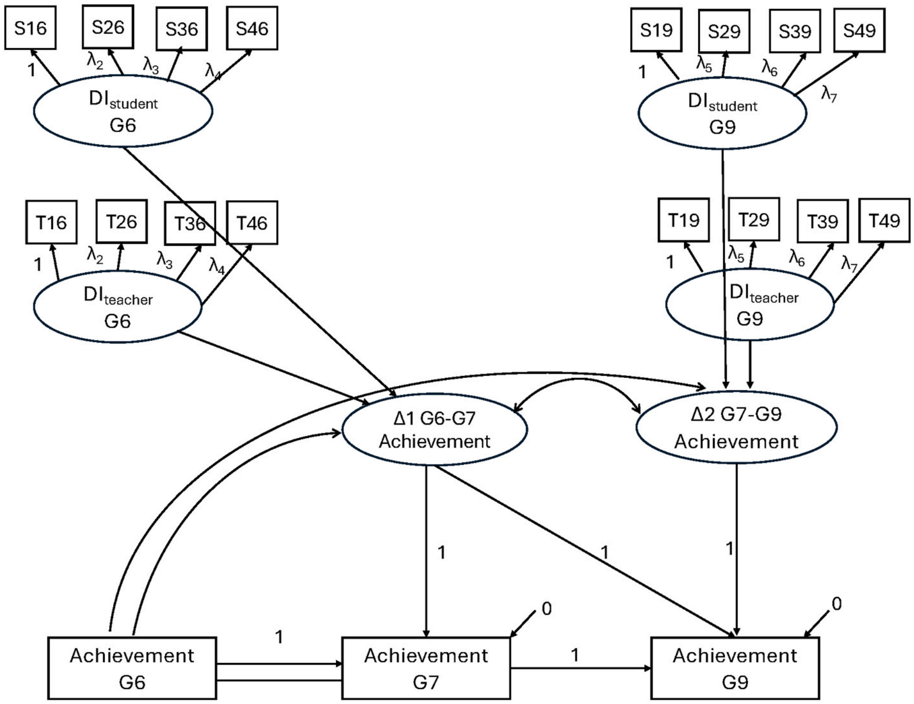

Fig. 1. Path diagram of latent neighbor change score model (example for achievement).

Table 1
Multilevel models for DI on academic achievement.

	Grade 6		Grade 9	
	Reading	Math	Reading	Math
Intercept	−0.71*	−0.50*	0.75*	1.51*
	(0.51)	(0.47)	(0.67)	(0.57)
Level 1 (within)				
DI, student report	0.02	−0.02	0.09*	0.06*
	(0.09)	(0.01)	(0.03)	(0.03)
Gender	−0.01	−0.17*	−0.01	−0.25*
	(0.02)	(0.01)	(0.03)	(0.03)
SEN-L	−0.14*	−0.19*	−0.10*	−0.16*
	(0.02)	(0.02)	(0.03)	(0.03)
Specific learning	−0.05*	−0.03*	−0.06	0.05
disability	(0.02)	(0.02)	(0.03)	(0.03)
Primary language	0.16*	0.10*	0.12*	0.03
	(0.02)	(0.02)	(0.04)	(0.03)
Educational background	0.08*	0.06*	0.80*	0.14*
	(0.02)	(0.02)	(0.03)	(0.03)
Basic cognitive skills	0.38*	0.50*	0.21*	0.44*
	(0.02)	(0.02)	(0.04)	(0.04)
Grade German	−0.19*	−0.09*	−0.16*	−0.15*
	(0.03)	(0.02)	(0.05)	(0.04)
Grade Math	−0.03	−0.15*	−0.12*	−0.17*
	(0.02)	(0.02)	(0.04)	(0.04)
Level 2 (between)				
DI, teacher report	0.20*	−0.01	−0.11	−0.17
	(0.10)	(0.08)	(0.15)	(0.12)
School type: multi-track	−0.71*	−0.69*	0.09	0.21*
(Ref. one-track)	(0.10)	(0.09)	(0.011)	(0.10)
School type: gymnasium	−0.72*	−0.62*	0.47*	0.58*
(Ref. one-track)	(0.09)	(0.08)	(0.09)	(0.07)
R^2 within	0.24	0.36	0.13	0.35
R^2 between	0.24	0.21	0.21	0.33

Note. $N_{\text{Grade 6}} = 4654$. $N_{\text{Grade 9}} = 1157$. Standardized coefficients; standard errors in parentheses. * = $p < .05$.

associations were not confounded by demographic or ability-related factors, mechanisms of non-participation or drop-out.

To investigate longitudinal effects (RQ2), we employed latent neighbor change models (McArdle, 2009), a longitudinal structural equation modelling approach that captures interindividual differences in intraindividual change between two or more measurement occasions. These models estimate latent change scores (Δ) between two consecutive measurement occasions, thereby enabling the analysis of increases or decreases in students' outcomes over time (Steyer et al., 1997). Our longitudinal models included two change parameters: one representing change from Grade 6 to Grade 7 (Δ_1) and another representing change from Grade 7 to Grade 9 (Δ_2). To examine how DI was related to these changes, we included teacher and student DI reports as predictors of latent change scores. DI reports of Grade 6 were used to predict changes between Grades 6 and 7, while DI reports of Grade 9 were used to predict changes from Grades 7 to 9 (example for achievement: Fig. 1).

To explore differential effects of DI for students with and without SEN-L (RQ3), we modeled an interaction effect. To account for the nested data structure, we relied on a complex design specification in M*plus* that adjusts standard errors for clustering (McNeish et al., 2017) by using a sandwich estimator.

In all models, subjective well-being was modeled as a latent three-factor construct, and reading and mathematics achievement as manifest variables. DI measures were specified as latent constructs, while covariates were treated as manifest variables.

4.4. *Missing data and robustness checks*

The data collection was affected by the Covid-19 pandemic. During the temporary school closures (spring term 2020), access to educational institutions was restricted. These circumstances led to a reduction in the number of participating schools and, consequently, in the sample size of participating students. At T2 and T3, 1832 and 1030 students, respectively, participated in the study.

Table 2
Multilevel models for DI on subjective well-being.

	Grade 6			Grade 9		
	EMO	SOC	ASC	EMO	SOC	ASC
Intercept	3.00* (0.09)	3.48* (0.07)	3.79* (0.09)	2.94* (0.18)	3.04* (0.15)	3.67* (0.17)
Level 1 (within)						
DI German, student report	0.29* (0.03)	0.25* (0.03)	0.17* (0.03)	0.40* (0.04)	0.26* (0.05)	0.22* (0.06)
DI Math, student report	0.23* (0.02)	0.17* (0.03)	0.22* (0.03)	0.10* (0.04)	0.16* (0.05)	0.13* (0.05)
Gender	0.10* (0.02)	0.01 (0.02)	−0.18* (0.02)	0.40 (0.04)	−0.10* (0.04)	−0.13* (0.05)
SEN-L	0.03 (0.02)	−0.05* (0.02)	−0.05* (0.02)	0.02 (0.04)	−0.03 (0.04)	−0.09* (0.04)
Specific learning disability	0.03 (0.02)	−0.01 (0.02)	−0.07* (0.02)	−0.03 (0.03)	−0.02 (0.04)	−0.07* (0.04)
Primary language	−0.06* (0.02)	0.01 (0.02)	−0.07* (0.02)	−0.06 (0.04)	−0.02 (0.04)	−0.11* (0.05)
Educational background	0.05* (0.02)	0.04 (0.02)	0.10* (0.02)	0.05 (0.04)	0.03 (0.04)	0.05 (0.04)
Basic cognitive skills	0.03 (0.02)	−0.03 (0.02)	0.14* (0.02)	0.01 (0.04)	0.03 (0.04)	0.11* (0.05)
Grade German	−0.12* (0.02)	−0.10* (0.02)	−0.22* (0.03)	−0.08 (0.04)	−0.03 (0.06)	−0.19* (0.05)
Grade Math	−0.05* (0.02)	0.01 (0.03)	−0.32* (0.03)	−0.05 (0.04)	0.06 (0.05)	−0.22* (0.05)
Level 2 (between)						
DI German, teacher report	0.02 (0.02)	−0.02 (0.02)	−0.01 (0.02)	−0.01 (0.04)	0.06 (0.04)	−0.02 (0.03)
DI Math, teacher report	−0.02 (0.02)	−0.01 (0.02)	0.03 (0.02)	0.01 (0.04)	−0.06* (0.03)	0.06* (0.03)
School type: multi-track (Ref. one-track)	0.02 (0.02)	0.01 (0.02)	−0.01 (0.01)	0.03 (0.04)	−0.03 (0.04)	0.07* (0.03)
School type: gymnasium (Ref. one-track)	0.04* (0.02)	0.04* (0.01)	−0.01 (0.01)	−0.01 (0.04)	0.01 (0.04)	0.05 (0.03)
R^2 within	0.24	0.15	0.33	0.21	0.13	0.25
R^2 between	0.01	0.01	0.01	0.01	0.01	0.01

Note. $N_{\text{Grade 6}} = 4654$. $N_{\text{Grade 9}} = 1157$. Standardized coefficients; standard errors in parentheses. EMO = Emotional well-being, SCO = Social inclusion, ASC = Academic self-concept. * = $p < .05$.

Table 3
Latent neighbor change model for academic achievement predicted by DI.

	Reading			Mathematics		
	G6	Δ1 G6-G7	Δ2 G7-G9	G6	Δ1 G6-G7	Δ2 G7-G9
Intercept	−0.85* (0.10)	0.42* (0.03)	0.13* (0.04)	−0.87* (0.09)	0.67* (0.02)	0.69* (0.04)
DI, student report (German | Math)		0.01 (0.03)	0.11* (0.03)		0.04* (0.02)	0.08* (0.03)
DI, teacher report (German | Math)		−0.01 (0.04)	−0.01 (0.06)		−0.06 (0.03)	−0.08 (0.06)
Controls						
Gender	0.01 (0.01)			−0.16* (0.01)		
SEN-L	−0.13* (0.02)			−0.17* (0.01)		
Specific learning disability	−0.04* (0.01)			−0.01 (0.01)		
Primary language	0.14* (0.02)			0.10* (0.01)		
Educational background	0.09* (0.02)			0.09* (0.01)		
Basic cognitive skills	0.73* (0.02)			0.50* (0.01)		
Grade German	−0.15* (0.02)			−0.08* (0.02)		
Grade Math	0.06* (0.02)			−0.16* (0.02)		
School type: multi-track (Ref. one-track)	0.01 (0.02)			0.03* (0.02)		
School type: gymnasium (Ref. one-track)	0.13* (0.02)			0.16* (0.02)		
R^2	0.22	0.03	0.25	0.37	0.03	0.27

Note. $N = 4654$. Standardized coefficients; standard errors in parentheses. DI = Differentiated instruction. * = $p < .05$.

It can be assumed that dropout and missing data were not systematically related to individual characteristics. Attrition analyses showed that few variables were significantly associated with non-participation, and differences were of negligible nature (Supplement; Tables B7–B8). Approximately 46% of students participated in at least two of the three measurement occasions (Table B8). Variables correlated with dropout and missingness at the individual level were included as control variables in the main analyses.

Sensitivity analyses excluding classes with fewer than three participating students (Model 6 in Tables B10–B13) showed that the results did not change substantially, indicating that the findings are robust with respect to class size at Level 2.

5. Results

5.1. Cross-sectional associations

The results of the multilevel regression for DI on students' academic achievement in Grades 6 and 9 are presented in Table 1. Student-reported DI showed small positive associations with reading and mathematics achievement only in Grade 9. Including control variables slightly reduced the regression coefficients of DI (Supplement, Tables B10 & B11). Teacher-reported DI was only positively associated with reading achievement in Grade 6.

Table 2 presents the results of the multilevel analyses for students' subjective well-being. Student-reported DI in both German and mathematics classes was weakly to moderately and positively correlated with subjective well-being across the three dimensions in both Grades 6 and 9. Including control variables caused only minimal changes to the estimates (Tables B12 & B13). In contrast, teacher-reported DI was not significantly associated with subjective well-being.

5.2. Longitudinal effects

Table 3 presents the results of the latent neighbor change models for academic achievement. In general, German and mathematics achievement increased over time. Student-reported DI positively predicted changes in mathematics achievement from Grades 6 to 7 and Grades 7 to 9, while changes in reading achievement were only found from Grades 7 to 9. All effect sizes were small-to-negligible. Teacher-reported DI

Table 4
Latent neighbor change model for subjective well-being predicted by DI.

	Emotional well-being			Social inclusion			Academic self-concept		
	G6	Δ1 G6-G7	Δ2 G7-G9	G 6	Δ1 G6-G7	Δ2 G7-G9	G 6	Δ1 G6-G7	Δ2 G7-G9
Intercept	3.16* (0.09)	−0.10* (0.03)	−0.33* (0.04)	3.51* (0.07)	0.09* (0.03)	−0.21* (0.05)	3.72* (0.07)	−0.11* (0.03)	0.01 (0.04)
Student report									
DI German		−0.13* (0.03)	0.30* (0.04)		−0.14* (0.03)	0.19* (0.05)		−0.12* (0.04)	0.13* (0.06)
DI Math		−0.10* (0.03)	0.06 (0.04)		−0.11* (0.03)	0.11* (0.05)		−0.15* (0.04)	0.11* (0.05)
Teacher report									
DI German		0.11* (0.05)	0.10 (0.13)		0.10 (0.05)	0.12 (0.15)		0.08 (0.06)	0.13 (0.13)
DI Math		−0.10 (0.06)	−0.24* (0.10)		−0.13* (0.07)	−0.31* (0.10)		−0.09 (0.07)	−0.20 (0.11)
Controls									
Gender	0.07* (0.01)			−0.02 (0.01)			−0.09* (0.01)		
SEN-L	0.02 (0.01)			−0.02 (0.01)			−0.03* (0.01)		
Specific learning disability	−0.001 (0.01)			−0.02 (0.01)			−0.06* (0.01)		
Primary language	−0.06* (0.01)			−0.01 (0.01)			−0.07* (0.01)		
Educational background	0.04* (0.01)			0.04* (0.01)			0.04* (0.01)		
Basic cognitive skills	0.02 (0.01)			−0.02 (0.01)			0.08* (0.01)		
Grade German	−0.09* (0.02)			−0.04* (0.01)			−0.13* (0.01)		
Grade Math	−0.06* (0.02)			−0.01 (0.01)			−0.19* (0.01)		
School type: multi-track (Ref. one-track)	0.01 (0.02)			0.01 (0.01)			0.01 (0.01)		
School type: gymnasium (Ref. one-track)	0.04* (0.01)			0.04* (0.02)			0.01 (0.01)		
R^2	0.01	0.05	0.16	0.02	0.06	0.16	0.08	0.05	0.08

Note. $N = 4654$. Standardized coefficients; standard errors in parentheses. Δ = change score variable. * = $p < .05$.

showed no significant effects on changes in academic achievement.

Regarding subjective well-being (Table 4), a general decrease was observed, except for social inclusion from Grades 6 to 7. Several significant effects emerged for DI from both perspectives. Student-reported DI in German classes was positively associated with changes in all three dimensions of subjective well-being from Grades 7 to 9. However, from Grades 6 to 7, this was negatively associated with changes in social inclusion. For student-reported DI in mathematics classes, small negative effects on changes in all three dimensions were found from Grades 6 to 7. Teacher-reported DI in German classes had a small positive effect on changes in emotional well-being and social inclusion from Grades 6 to 7, whereas teacher-reported DI in mathematics classes revealed small negative effects on changes in the same two dimensions from Grades 7 to 9.

5.3. Differential effects

The latent neighbor change models including interaction terms for DI and SEN-L showed two significant interaction effects. The interaction between student-reported DI and SEN-L negatively predicted change in mathematics achievement from Grades 7 to 9 (Table 5), indicating that the effect of DI on change in mathematics achievement was negative for students with SEN-L, but not for students without SEN-L (Fig. 2). In contrast, teacher-reported DI did not show any significant interaction effects on changes in academic achievement for either group.

Focusing on subjective well-being (Table 6), the interaction between student-reported DI in German classes and SEN-L positively predicted change in social inclusion from Grades 7 to 9, indicating a positive effect for students with SEN-L but not for other students (Fig. 3). Again, no significant interaction effects were found for teacher-reported DI.

6. Discussion

Our cross-sectional results indicate that DI is in some cases positively, yet weakly, associated with students' academic achievement and — particularly student-reported DI —with their subjective well-being both in Grades 6 and 9 (RQ1). Longitudinally, DI predicted some changes in students' academic achievement and subjective well-being during lower secondary education (RQ2), albeit inconsistently and with small-to-negligible effect sizes. These effects of DI on changes in students' academic achievement were predominantly positive, whereas several negative effects were observed for changes in their subjective well-being. Furthermore, the findings indicate two differential effects for students with SEN-L compared to other students (RQ3), although not necessarily in the expected directions.

6.1. Challenges of implementing DI

While the general pattern of our cross-sectional analyses aligns with previous studies—primarily conducted at the primary school level (Graham et al., 2021)—our findings add to prior research by investigating effects of DI at both earlier and later stages of lower secondary education. While the associations of DI with academic achievement were rather small and inconsistent across both grades, the positive associations of (student-reported) DI with subjective well-being were consistent. More importantly, our study extends research on DI by longitudinal analyses predicting changes in students' academic achievement and subjective well-being during lower secondary education, although effects were inconsistent and small to negligible.

A possible explanation for the contradictory results is that effectiveness of DI may depend on teachers' capacity to dynamically adjust

Table 5
Latent neighbor change model for academic achievement predicted by DI and interaction for SEN-L.

	Reading			Mathematics		
	G6	Δ1 G6-G7	Δ2 G7-G9	G6	Δ1 G6-G7	Δ2 G7-G9
Intercept	−1.37* (0.13)	0.55* (0.05)	0.11* (0.06)	−1.37* (0.12)	0.71* (0.04)	0.47* (0.04)
DI, student report		−0.02 (0.07)	0.28* (0.13)		0.10 (0.06)	−0.01 (0.05)
DI, teacher report		−0.03 (0.15)	0.05 (0.16)		−0.01 (0.10)	−0.10 (0.10)
SEN-L	−0.72* (0.08)	0.35* (0.10)	−0.05 (0.12)	−0.76* (0.07)	0.17* (0.09)	−0.19* (0.07)
DI student report x SEN-L		−0.06 (0.14)	0.13 (0.25)		0.10 (0.11)	−0.29* (0.10)
DI teacher report x SEN-L		0.01 (0.25)	0.20 (0.0)		0.33 (0.23)	0.12 (0.21)
Controls						
Gender	0.02 (0.03)			−0.37* (0.03)		
Specific learning disability	−0.13* (0.04)			−0.04 (0.04)		
Primary language	0.51* (0.06)			0.34* (0.05)		
Educational background	0.05* (0.01)			0.04* (0.01)		
Basic cognitive skills	0.13* (0.01)			0.17* (0.01)		
Grade German	−0.20* (0.03)			−0.10* (0.02)		
Grade Math	0.08* (0.02)			−0.18* (0.02)		
School type: multi-track (Ref. one-track)	0.02 (0.05)			0.08* (0.04)		
School type: gymnasium (Ref. one-track)	0.69* (0.10)			0.80* (0.07)		

Note. N = 4654. Unstandardized coefficients; standard errors in parentheses. * = $p < .05$.

instructional guidance. While teachers may implement DI, it cannot be assumed that their instructional decisions are consistently informed by students' learning readiness or supported by relevant data (Karst et al., 2022). Although such responsiveness represents a crucial prerequisite for effective DI (Tomlinson et al., 2003), formative assessment (i.e., data-based decision-making) appears to be infrequently utilized by teachers in classroom practice (Jager et al., 2025). This limited use may reflect the substantial cognitive and organizational demands associated with DI, which requires a high level of professional competence and pedagogical expertise (Brühwiler & Vogt, 2020). In this vein, our findings regarding the negative effects of DI on changes in students' subjective well-being – especially on social inclusion – could indicate teachers' misconceptions about DI or an overly narrow understanding of it (Dack, 2019). Teachers may primarily rely on ability grouping as an organizational format to implement differentiation (Prast et al., 2018), which in turn may negatively affect students' well-being and social inclusion in class.

Given the pronounced heterogeneity of student populations in inclusive classrooms, it is plausible that teachers in such settings may not be sufficiently prepared to address each individual student's specific needs (Goldan et al., 2022). Although DI plays a pivotal role for the successful implementation of inclusive education (Langelaan et al., 2024), its implementation has not yet reached its full potential. These limitations may reflect broader systemic challenges, including limited resources, time constraints, and inadequate teacher training (van Geel et al., 2018). The challenges of implementing DI may be further amplified in secondary education because subject teachers have less instructional time with each class and face greater organizational complexity.

6.2. *Different perspectives on DI*

Overall, there was a noticeable discrepancy in the results depending on whether DI was assessed from the students' or teachers' perspectives. Several factors may explain this divergence in findings, which was somewhat more pronounced for students' subjective well-being than for achievement. One potential explanation for these contradictory findings is that students' well-being may improve when they perceive their teachers as proactively addressing their needs. Teachers, in contrast, may be less sensitive to students' subjective well-being and may not explicitly consider it when implementing DI practices (Zurbriggen et al., 2025).

Teachers' perceptions of their instructional practices may also be influenced by self-presentation bias, whereas students' evaluations may be shaped by factors such as teacher popularity or personal preferences (Tetzlaff et al., 2022). However, it would also be plausible to assume that both perspectives used to describe DI capture related but partially distinct constructs. Consistent with this, previous research has shown that the level of agreement between student and teacher ratings of DI is generally low to moderate (Kunter & Baumert, 2006; Tetzlaff et al., 2022).

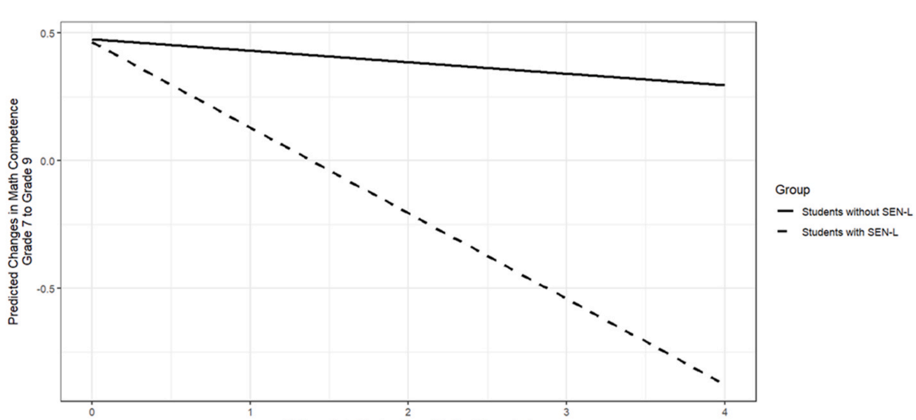

Fig. 2. Interaction effect for mathematics achievement.

Table 6
Latent neighbor change model for subjective well-being predicted by DI and interaction for SEN-L.

	Emotional well-being			Social inclusion			Academic self-concept		
	G6	Δ1 G6-G7	Δ2 G7-G9	G 6	Δ1 G6-G7	Δ2 G7-G9	G 6	Δ1 G6-G7	Δ2 G7-G9
Intercept	2.95* (0.09)	−0.08* (0.03)	−0.23* (0.05)	3.22* (0.09)	0.06* (0.03)	−0.10* (0.04)	3.86* (0.08)	−0.05* (0.03)	0.01 (0.04)
Student report									
DI German		−0.03 (0.17)	0.49 (0.22)		−0.13 (0.08)	0.37* (0.11)		−0.02 (0.14)	0.25 (0.16)
DI Math		−0.18 (0.13)	0.09 (0.11)		−0.06 (0.09)	0.08 (0.06)		−0.12 (0.13)	0.01 (0.09)
SEN-L	0.10 (0.05)	−0.03 (0.07)	−0.02 (0.10)	−0.07 (0.034)	0.03 (0.06)	−0.01 (0.07)	−0.12* (0.04)	0.01 (0.06)	0.03 (0.08)
DI German x SEN-L		0.18 (0.33)	0.17 (0.43)		−0.15 (0.17)	0.38* (0.19)		0.03 (0.29)	0.19 (0.32)
DI Math x SEN-L		−0.23 (0.26)	0.06 (0.22)		−0.03 (0.17)	−0.01 (0.13)		−0.11 (0.25)	−0.19 (0.18)
Teacher report									
DI German		0.26 (0.28)	0.23 (0.25)		0.09 (0.15)	0.08 (0.17)		0.17 (0.31)	−0.04 (0.20)
DI Math		0.00 (0.26)	−0.56 (0.39)		−0.05 (0.17)	−0.19 (0.29)		−0.09 (0.29)	−0.24 (0.20)
DI German x SEN-L		0.12 (0.52)	0.17 (0.42)		−0.13 (0.29)	−0.05 (0.29)		0.13 (0.58)	−0.39 (0.37)
DI Math x SEN-L		0.42 (0.52)	−0.46 (0.69)		0.31 (0.32)	0.29 (0.60)		0.05 (0.56)	−0.15 (0.56)
Controls									
Gender	0.12* (0.03)			−0.03 (0.02)			−0.19* (0.02)		
Specific learning disability	0.001 (0.04)			−0.03 (0.02)			−0.14* (0.03)		
Primary language	−0.18* (0.04)			−0.03 (0.03)			−0.20* (0.03)		
Educational background	0.02* (0.01)			0.01* (0.004)			0.02* (0.01)		
Basic cognitive skills	0.01 (0.01)			−0.01 (0.01)			0.02* (0.01)		
Grade German	−0.10* (0.02)			−0.04* (0.01)			−0.14* (0.03)		
Grade Math	−0.06* (0.02)			−0.01 (0.01)			−0.19* (0.01)		
School type: multi-track (Ref. one-track)	0.01 (0.03)			0.01 (0.03)			−0.001 (0.03)		
School type: Gymnasium (Ref. one-track)	0.16* (0.05)			0.14* (0.04)			0.05 (0.04)		

Note. $N = 4654$. Unstandardized coefficients; standard errors in parentheses. Δ = change score variable. * = $p < .05$.

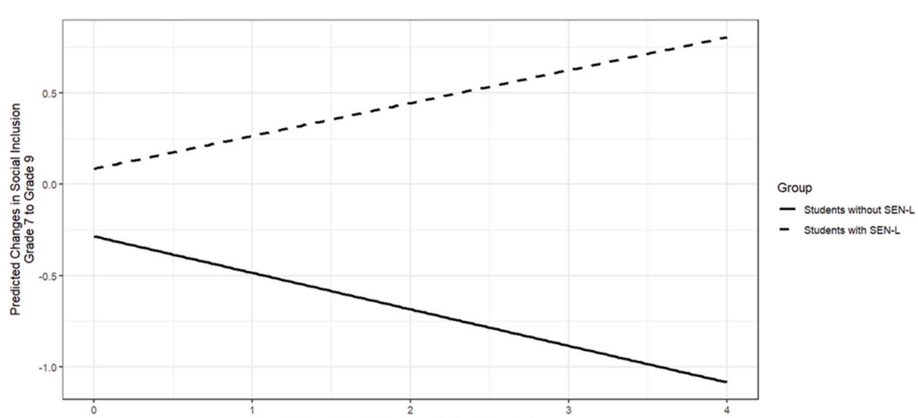

Fig. 3. Interaction effect for social inclusion.

6.3. (Un)expected differential effects

Only two significant interaction effects were found for students with versus without SEN-L. As anticipated, findings are mixed. First, our findings indicate that adapting instructions may support the self-perceived social inclusion of students with SEN-L in class. Given that students with SEN are often less included socially, the potential of DI in this regard should be further explored. However, no differential effects were found for the other components of subjective well-being, which can be interpreted as a sign for equality of conditions (Valiandes, 2015).

Second, increases in mathematics achievement were negatively predicted by (student-reported) DI for students with SEN-L in higher

Grades, but not for students without SEN-L. Unlike most previous research on differential effects of DI on academic achievement—which primarily compared students with low versus high academic achievement—our focus was on students with SEN-L versus their peers. Although students with SEN-L typically demonstrate lower achievement, substantial differences exist between these groups (Gehrer & Nusser, 2020), limiting comparability of our findings with other studies. Nevertheless, the negative interaction effect for mathematics for students with SEN-L partly contradicts theoretical assumptions of an aptitude-treatment interaction. According to the expertise reversal effect—one prominent example of such an interaction effect—negative effects would be expected for students with moderate to high achievement compared to those with lower achievement (Tetzlaff et al., 2025). That said, how DI can achieve educational equity in the sense of adequacy for all students remains a subject of considerable debate (Dumont & Ready, 2023).

6.4. Limitations and future directions

Several limitations should be acknowledged. First, DI was assessed from student and teacher perspectives using two distinct scales with slight differences in item wording. Thus, findings from these two perspectives are not directly comparable, despite the intentional theoretical alignment of item content. Future studies might consider evaluating consistency between student and teacher reports of DI or including expert classroom observations (Kunter & Baumert, 2006). Second, although DI was conceptualized as a comprehensive construct in line with Tomlinson's (2017) framework, the items used in this study captured only specific aspects of the DI dimensions. Thus, findings are not transferable to all DI practices. Future research would benefit from incorporating a broader range of differentiation practices and examining DI as a multidimensional construct.

Third, student-reported DI was not assessed in Grade 7, which resulted in uneven intervals and restricted the scope of the longitudinal analyses. This major limitation may partly explain the absence of longitudinal effects of DI on change in students' outcomes from Grades 6 to 7. Fourth, as the sample was drawn from inclusive schools in Germany, our findings primarily reflect the German educational context. Fifth, data collection during Grade 6 was affected by the Covid-19 pandemic, and new digital teaching approaches may have influenced student outcomes. The negative or absent effects during this period may therefore partly reflect pandemic-related disruptions. Therefore, future studies are needed to replicate the findings while controlling for other potential confounding variables at the instructional level.

Finally, since our study relied on a non-experimental design, differential effects for students with SEN-L compared to those without SEN-L cannot be interpreted as "treatment effects" or causal effects. Future studies could examine the expertise reversal effect (Tetzlaff et al., 2025) in the context of inclusive education or focus on other student groups (e.g., distinct behavioral profiles) or other socio-emotional outcomes, such as learning enjoyment.

6.5. Conclusion

Despite some limitations, this study extends research on DI by focusing on both students' academic achievement and subjective well-being in lower secondary education, and by identifying areas that warrant further investigation. While cross-sectional analyses suggested several positive associations between DI and student outcomes, longitudinal results were more inconsistent, including small to negligible positive but also some negative effects. Student-reported DI emerged as a more consistent predictor of both outcomes than teacher-reported DI, underscoring the importance of incorporating both perspectives when investigating adaptive instructional practices. The two small differential effects observed for students with SEN-L compared to other students may indicate a need for more targeted DI approaches. Overall, the findings of the present study contribute to the growing body of evidence on DI, while calling for continued research to clarify its underlying mechanisms and to improve its effective implementation in inclusive classrooms, addressing the diverse needs of all students.

Funding

This study was supported by the German Federal Ministry of Education and Research (BMBF) under grants IN1503A, IN1503B, IN1503C, and IN1503D.

CRediT authorship contribution statement

Carmen L.A. Zurbriggen: Conceptualization, Methodology, Writing – original draft. **Marcela Pozas:** Writing – original draft, Writing – review & editing. **Karin Gehrer:** Conceptualization, Resources, Writing – review & editing. **Lena Nusser:** Conceptualization, Formal analysis, Methodology, Writing – review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2026.102431.

Data availability

Data is available at:https://www.lifbi.de/en-us/Start/Data-Services/Data-and-Documentation/SUF

References

Alnahdi, G. H., Lindner, K. T., & Schwab, S. (2022). Teachers' implementation of inclusive teaching practices as a potential predictor for students' perception of academic, social and emotional inclusion. *Frontiers in Psychology, 13*, Article 917676. https://doi.org/10.3389/fpsyg.2022.917676

Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M., & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality, 74*, 83–94. https://doi.org/10.1016/j.jrp.2018.02.007

Brauns, H., Scherer, S., & Steinmann, S. (2003). The CASMIN educational classification in international comparative research. In J. Hoffmeyer-Zlotnik, & C. Wolf (Eds.), *Advances in cross-national comparison* (pp. 221–244). Springer. https://doi.org/10.1007/978-1-4419-9186-7_11.

Brühwiler, C., & Vogt, F. (2020). Adaptive teaching competency: Effects on quality of instruction and learning outcomes. *Journal for Educational Research Online, 12*(1), 119–142. https://doi.org/10.18747/phsg-coll3/id/1213

Cronbach, L. J., & Snow, R. E. (1981). *Aptitudes and instructional methods: A handbook for research on interactions*. Ardent Media.

Dack, H. (2019). Understanding teacher candidate misconceptions and concerns about differentiated instruction. *The Teacher Educator, 54*(1), 22–45. https://doi.org/10.1080/08878730.2018.1485802

Datnow, A., & Hubbard, L. A. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record: The Voice of Scholarship in Education, 117*(4), 1–26. https://doi.org/10.1177/0161468115117004

Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review, 24*, 31–54. https://doi.org/10.1016/j.edurev.2018.02.002

Dumont, H., & Ready, D. D. (2023). On the promise of personalized learning for educational equity. *npj Science of Learning, 8*, 26. https://doi.org/10.1038/s41539-023-00174-x

Förster, N., Kawohl, E., & Souvignier, E. (2018). Short- and long-term effects of assessment-based differentiated reading instruction in general education on reading fluency and reading comprehension. *Learning and Instruction, 56*, 98–109. https://doi.org/10.1016/j.learninstruc.2018.04.009

Gebhardt, M., Schwab, S., Krammer, M., Gasteiger-Klicpera, B., & Sälzer, C. (2014). Erfassung von individualisiertem Unterricht in der Sekundarstufe I [Measuring teaching practices in inclusive settings in secondary classrooms]. *Zeitschrift für Bildungsforschung, 4*, 303–316. https://doi.org/10.1007/s35834-014-0095-7

Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online, 5*(2), 50–79.

Gehrer, K., & Nusser, L. (2020). Binnendifferenzierender Deutschunterricht und dessen Einfluss auf die Lesekompetenzentwicklung in der Sekundarstufe I [Adaptive instruction in German classes and its effect on reading competence development in secondary school]. *Journal for Educational Research Online, 12*(2), 166–189. https://doi.org/10.25656/01:20976

Goldan, J., Nusser, L., & Gebel, M. (2022). School-related subjective well-being of children with and without special educational needs in inclusive classrooms. *Child Indicators Research, 15*, 1313–1337. https://doi.org/10.1007/s12187-022-09914-8
Graham, L. J., de Bruin, K., Lassig, C., & Spandagou, I. (2021). A scoping review of 20 years of research on differentiation: Investigating conceptualisation, characteristics, and methods used. *The Review of Education, 9*(1), 161–198. https://doi.org/10.1002/rev3.3238
Hascher, T. (2012). Well-being and learning in school. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3453–3456). Springer. https://doi.org/10.1007/978-1-4419-1428-6_1832.
Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
Jager, L., Denessen, E., Cillessen, A. H., & Meijer, P. C. (2025). Instructional differentiation in secondary education: Teachers' actions and reasoning. *Teaching and Teacher Education, 159*. https://doi.org/10.1016/j.tate.2025.105009
Karst, K., Bonefeld, M., Dotzel, S., Fehringer, B. C. O. F., & Steinwascher, M. (2022). Data-based differentiated instruction: The impact of standardized assessment and aligned teaching material on students' reading comprehension. *Learning and Instruction, 79*, Article 101597. https://doi.org/10.1016/j.learninstruc.2022.101597
Kulakow, S. (2020). Academic self-concept and achievement motivation among adolescent students in different learning environments: Does competence-support matter? *Learning and Motivation, 70*, Article 101632. https://doi.org/10.1016/j.lmot.2020.101632
Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231–251. https://doi.org/10.1007/s10984-006-9015-7
Langelaan, B. N., Gaikhorst, L., Smets, W., & Oostdam, R. J. (2024). Differentiating instruction: Understanding the key elements for successful teacher preparation and development. *Teaching and Teacher Education, 140*, Article 104464. https://doi.org/10.1016/j.tate.2023.104464
Lindner, K. T., Schwab, S., Emara, M., & Avramidis, E. (2023). Do teachers favor the inclusion of all students? A systematic review of primary schoolteachers' attitudes towards inclusive education. *European Journal of Special Needs Education, 38*(6), 766–787. https://doi.org/10.1080/08856257.2023.2172894
McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*(1), 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612
McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114–140. https://doi.org/10.1037/met0000078
Muthén, B. O., & Muthén, L. K. (2017). *Mplus (Version 8)*. Muthén & Muthén.
Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E., & Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for Educational Research Online, 5*(2), 80–109.
Nusser, L., & Gehrer, K. (2020). Addressing heterogeneity in secondary education: Who benefits from differentiated instruction in German classes? *International Journal of Inclusive Education, 29*(12), 2220–2237. https://doi.org/10.1080/13603116.2020.1862407
O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181–206. https://doi.org/10.1207/s15326985ep4103_4
Obrovská, J., Svojanovský, P., & Sharma, U. (2025). Teachers' inclusive core practices and all students' perspectives. *Teaching and Teacher Education, 165*, Article 105120. https://doi.org/10.1016/j.tate.2025.105120
Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research, 88*(2), 205–242. https://doi.org/10.3102/0034654317743198
Pozas, M., Letzel, V., Lindner, K., & Schwab, S. (2021). DI (differentiated instruction) does matter! the effects of DI on secondary school students' well-being, social inclusion and academic self-concept. *Frontiers in Education, 6*, Article 729027. https://doi.org/10.3389/feduc.2021.729027
Pozas, M., Letzel, V., & Schneider, C. (2019). Teachers and differentiated instruction: Exploring differentiation practices to address student heterogeneity. *Journal of Research in Special Educational Needs, 20*(3), 217–230. https://doi.org/10.1111/1471-3802.12481
Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM - Mathematics Education, 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4
Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. (2015). Readiness-based differentiation in primary school mathematics: Expert recommendations and teacher self-assessment. *Frontline Learning Research, 3*(2), 90–116. https://doi.org/10.14786/flr.v3i2.163
Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2018). Differentiated instruction in primary mathematics: Effects of teacher professional development on student achievement. *Learning and Instruction, 54*, 22–34. https://doi.org/10.1016/j.learninstruc.2018.01.009
Roy, A., Guay, F., & Valois, P. (2013). Teaching to address diverse learning needs: Development and validation of a differentiated instruction scale. *International Journal of Inclusive Education, 17*(11), 1186–1204. https://doi.org/10.1080/13603116.2012.743604
Roy, A., Guay, F., & Valois, P. (2015). The big-fish–little-pond effect on academic self-concept: The moderating role of differentiated instruction and individual achievement. *Learning and Individual Differences, 42*, 110–116. https://doi.org/10.1016/j.lindif.2015.07.009
Saloviita, T. (2018). How common are inclusive educational practices among Finnish teachers. *International Journal of Inclusive Education, 22*, 560–575. https://doi.org/10.1080/13603116.2017.1390001
Schmitt, M., Gresch, C., & Labsch, A. (2023). Challenges in sampling students with and without special educational needs in inclusive settings for an educational longitudinal study. *Survey Methods, 1*(2). https://doi.org/10.13094/SMIF-2023-00011
Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*, 499–545. https://doi.org/10.3102/0034654307310317
Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology, 10*, 2366. https://doi.org/10.3389/fpsyg.2019.02366
Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research, 86*(4), 849–899. https://doi.org/10.3102/0034654316675417
Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online, 2*, 21–33.
Tetzlaff, L., Hartmann, U., Dumont, H., & Brod, G. (2022). Assessing individualized instruction in the classroom: Comparing teacher, student, and observer perspectives. *Learning and Instruction, 82*, 1–10. https://doi.org/10.1016/j.learninstruc.2022.101655
Tetzlaff, L., Simonsmeier, B., Peters, T., & Brod, G. (2025). A cornerstone of adaptivity – A meta-analysis of the expertise reversal effect. *Learning and Instruction, 98*. https://doi.org/10.1016/j.learninstruc.2025.102142
Theisel, A. (2014). Ein Faktorenmodell zu Qualitätsmerkmalen des Unterrichts mit sprachbeeinträchtigten Kindern [A factor model on quality characteristics of teaching children with language impairments]. *Forschung Sprache, 2*(2), 36–57.
Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners*. ASCD.
Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms*. ASCD.
Tomlinson, C. A., Brigthon, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted, 27*(2–3), 119–145. https://doi.org/10.1177/01623532030270
Valiandes, S. (2015). Evaluating the impact of differentiated instruction on literacy and reading in mixed ability classrooms: Quality and equity dimensions of education effectiveness. *Studies In Educational Evaluation, 45*, 17–26. https://doi.org/10.1016/j.stueduc.2015.02.005
van der Scheer, E. A., & Visscher, A. J. (2018). Effects of a data-based decision-making intervention for teachers on students' mathematical achievement. *Journal of Teacher Education, 69*(3), 307–320. https://doi.org/10.1177/0022487117704170
van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2018). Capturing the complexity of differentiated instruction. *School Effectiveness and School Improvement, 30*(1), 51–67. https://doi.org/10.1080/09243453.2018.1539013
Venetz, M., Zurbriggen, C., Eckhart, M., Schwab, S., & Hessels, M. (2015). The perceptions of inclusion questionnaire (PIQ) (German version) https://piqinfo.ch.
Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427–450. https://doi.org/10.1007/BF02294627
Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., Carstensen, C. H., & Lockl, K. (2019). Development of competencies across the life course. In H.-P. Blossfeld, & H.-G. Roßbach (Eds.), *Education as a lifelong process, 3* pp. 57–81). Fachmedien Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-23162-0_4.
Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. *Jahrgangsstufe [Berlin test to assess fluid and crystallized intelligence for grades 8 to 10]. Hogrefe*.
Ziernwald, L., Hillmayr, D., & Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms—a systematic review. *Journal of Advanced Academics, 33*(4), 540–573. https://doi.org/10.1177/1932202x221112931
Zurbriggen, C., Schwab, S., Kalkusch, I., Neuhauser, A., Lanfranchi, A., & Klaver, P. (2025). Explaining teachers' judgement accuracy of students' subjective well-being by students' individual characteristics and the teacher-student relationship. *Learning and Individual Differences, 122*, Article 102754. https://doi.org/10.1016/j.lindif.2025.102754
Zurbriggen, C. L. A., Venetz, M., Schwab, S., & Hessels, M. G. P. (2019). A psychometric analysis of the student version of the perceptions of inclusion questionnaire (PIQ). *European Journal of Psychological Assessment, 35*(5), 641–649. https://doi.org/10.1027/1015-5759/a000443

Carmen Zurbriggen is a full professor in Special Education at the University of Fribourg, Switzerland. Lena Nusser is head of the research unit Learning Environment and School Education at the Leibniz Institute for Educational Trajectories (LIfBi), Bamberg, Germany. Marcela Pozas is full professor in Educational Diversity Management at the University of Luxembourg. Karin Gehrer is post doctoral researcher at the LIfBi. Their research focuses on inclusive education, learning and subjective well-being in diverse school contexts.