

Promoting High-Achieving Students Through Differentiated Instruction in Mixed-Ability Classrooms—A Systematic Review

Journal of Advanced Academics
2022, Vol. 33(4) 540–573
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1932202X221112931
journals.sagepub.com/home/joa



Lisa Ziernwald¹ , Delia Hillmayr¹,
and Doris Holzberger¹

Abstract

Promoting high-achieving students plays an important role in the school context. Hence, one promising support measure within the mixed-ability classroom is differentiated instruction (DI). The current systematic review examined (1) the impact of DI on high-achieving students' outcomes, (2) to what extent DI is used, (3) how useful teachers and high-achieving students perceive DI, and (4) which barriers and facilitators are encountered in DI's implementation. Forty-nine studies from 2000 to 2019 were included. Differentiated instruction impacted high-achieving students' academic achievement and motivational-affective characteristics predominantly positive. However, there was considerable heterogeneity between and within studies. Teachers typically did not use DI for high-achieving students proactively nor on a regular basis. However, teachers and high-achieving students perceived DI as valuable for encouraging high-achieving students. The barriers found might help to explain discrepancies between the extent of usage and the perceived utility, whereas the identified facilitators suggest how to overcome these barriers.

Keywords

gifted, differentiated instruction, high-achieving students, mixed-ability, classroom, systematic mixed-methods review, talent development

¹Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Department of Educational Sciences, Technical University of Munich, Munich, Germany

Corresponding Author:

Lisa Ziernwald, Technical University of Munich, Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Department of Educational Sciences, Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany.

Email: lisa.ziernwald@tum.de

Encouraging all students according to their individual skills and helping them realize their full potential is essential in the context of educational equity (General Assembly of the United Nations, 2019; Kultusministerkonferenz, 2009, 2015). High-achieving students are thereby described as a crucial resource for society and the world (Kahveci & Akgül, 2014; Ninkov, 2020). The current research synthesis investigated the promotion of high-achieving students inside mixed-ability classrooms.

The focus on mixed-ability classrooms has various advantages. First, most high-achieving students are taught most of the time in mixed-ability classrooms (Betts, 2004; Dreeszen, 2009; White et al., 2003; Ysseldyke et al., 2004; Yuen et al., 2018). Second, special support measures outside the mixed-ability classroom have decreased (Latz et al., 2008; Tomlinson et al., 2003) and are frequently not accessible because of limited financial and organizational resources (Endepohls-Ulpe, 2017). Third, compared to pull-out programs, where students are removed from their regular classroom to supplement their classroom curriculum for a specific amount of time per week (Gubbins, 2013), students can be challenged daily through promotion inside mixed-ability classrooms (Dreeszen, 2009). Fourth, the inclusion of high-achieving students in mixed-ability classrooms can positively affect them, as well as their classmates with differing abilities, by raising the standard of learning (Ninkov, 2020). Fifth, high-achieving students, their parents, and teachers favor support measures integrated into the mixed-ability classroom compared to segregated measures like special classes (Shayshon et al., 2014; Sparfeldt et al., 2004).

However, so far, only little is known about the effectiveness of instructional practices for high-achieving students in mixed-ability classrooms (Barbier et al., 2022). One promising approach for promoting high-achieving students within mixed-ability classrooms is differentiated instruction (DI). It addresses students' heterogeneity and aims to educate them adequately according to their individual needs (Tomlinson, 2001).

Theoretical Framework of DI and Its Role for High-Achieving Students

Differentiated instruction can be seen as a part of the broader construct *differentiation*, which not only includes DI during a lesson but also student assessment, evaluation, philosophical aspects, and more general principles (cf. Smale-Jacobse et al., 2019; Tomlinson, 2014). To attain a clear focus despite the fuzzy construct of differentiation (Deunk et al., 2018), we focused the current systematic review on DI in mixed-ability classrooms, which contains two main aspects: (a) instructional practices and (b) organizational aspects (Smale-Jacobse et al., 2019). DI covers various instructional practices such as tiered tasks, scaffolds, and adaptive learning materials. The organizational aspect contains structures for DI, such as within-class grouping arrangements (Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019; Tomlinson, 2014). Another construct that considers the needs of individual students and provides corresponding learning opportunities for each student is individualization. In the

current systematic review, individualization is included as a part of DI as it can be seen as an extreme type of DI, where instructional practices are not adapted for student sub-groups (e.g., high-achievers) but for the individual student (Bohl et al., 2012; Dumont, 2018; Smale-Jacobse et al., 2019).

Three current research syntheses (Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019) and a literature review (Bondie et al., 2019) focused on the definition, operationalization, barriers and facilitators, and effectiveness of DI and underlined its potential to promote students according to their individual needs. Although DI is integrated into different models for the education of high-achieving students (e.g., Kaplan, 1986; Maker & Schiever, 2010; Moon et al., 2009; Williams, 1986), DI has not been investigated through a research synthesis with the focus on high-achieving students in mixed-ability classrooms so far.

The overarching aim of DI is to build on the commonalities of the whole class while also considering students' differences so that everyone has the opportunity to learn effectively (Tomlinson, 2001). Therefore, DI can, by definition, improve the learning of all students in mixed-ability classrooms but might be particularly helpful for high-achieving students (Troxclair, 2000; Yuen et al., 2018). Borland (2005, pp. 1–2) stated that “high-achieving or high-ability students are among those who are the most ill-served when curriculum and instruction are not differentiated.”

Research on the Impact of DI on Students' Outcomes

Recent studies have addressed the impact of DI in different contexts. Bal (2016), as an example, investigated DI through tiered instructional practices, including appropriately differentiated worksheets, activities, and other materials, according to the students' readiness level and learning styles in mathematics lessons. It was shown that these sixth-grade students improved their mathematics achievement compared to a control group taught traditionally. The students' positive opinions regarding their mathematics achievement and motivational-affective development supported these objective test results. In mixed-ability classrooms, DI can be either evaluated concerning the general impact (i.e., on all students, cf. Bal, 2016) or concerning differential impacts (i.e., on low-, average-, and high-achieving students, cf. Meyer et al., 2011).

Moreover, the impact on student outcomes can be evaluated through academic achievement and motivational-affective characteristics. Motivational-affective characteristics, such as attitudes toward reading, performance avoidance, and motivational beliefs, are related to academic achievement and play an important role in the academic and occupational success (Gutman & Schoon, 2013; OECD, 2013, 2019). This perspective is also reflected in the field of giftedness, where multidimensional approaches—including achievement factors *and* motivational-affective characteristics—are strongly represented (Heller et al., 2005; Stoeger, 2009). Previous research syntheses showed that enrichment and acceleration for high-achieving students could have differential effects on academic achievement and motivational-affective characteristics (Kim, 2016; Steenbergen-Hu & Moon, 2011). The current study summarized different motivational and emotional constructs, sometimes also named

noncognitive characteristics (Heller et al., 2005), such as motivational beliefs and attitudes under the term *motivational-affective characteristics*.

Recently, increased efforts have been made to synthesize the existing empirical evidence regarding the impact of DI on all students—-independent of whether they were identified as high achievers or not. Two current research syntheses investigated DI in the context of primary (Deunk et al., 2018) and secondary education (Smale-Jacobse et al., 2019). Deunk et al. (2018) investigated the effects of four categories of DI, namely between-class and within-class homogeneous ability grouping, computerized differentiation tools, and differentiation as part of a broader program on students' language or mathematics achievement. They concluded that DI—either based on computer systems or embedded in a broader program—showed small to moderate positive effects on students' achievement for the whole student sample. Smale-Jacobse et al. (2019) investigated the effectiveness of within-class DI from 2006 to 2016 on students' academic achievement. They concluded that many different forms of DI (e.g., ability grouping, tiering, and individualization) had been used in the primary studies and that DI showed mostly small to moderate positive effects on students' academic achievement. Moreover, Smale-Jacobse et al. (2019) emphasized that future studies on DI should investigate academic achievement along with motivational-affective characteristics. A systematic review and meta-analysis investigated DI in the context of literacy instruction (Puzio et al., 2020) and found an overall positive effect on students' literacy achievement. Again, various forms of DI (e.g., individualization, alternate curriculum) were investigated. Although these research syntheses showed that DI could positively influence students' academic achievement in the whole investigated mixed-ability sample, a systematic investigation regarding the impact of DI on high-achieving students in mixed-ability classrooms did not take place so far.

Conditions for Implementing DI

To promote high-achieving students, the impact of DI is important, and that teachers *regularly implement* this approach in their mixed-ability classrooms. Without implementing DI, the potential that may exist cannot unfold. If, for example, hands-on activities were used every day or once a week, students' science achievement was significantly higher compared to students who were engaged in these activities only once a month or less (Stohr-Hunt, 1996). Reis and Renzulli (2010) reviewed the literature on gifted education and concluded that, at least in America, DI has only been implemented to a very small extent. This is in accordance with studies from the 1990s that found that DI is not used regularly for high-achieving students (Archambault et al., 1993; Matthews, 1992; Westberg et al., 1993). Westberg et al. (1993) observed high-achieving students' instruction in 46 grades 3 and 4 mixed-ability classrooms. It was found that 84% of the high-achieving students' instructional activities across five subjects included no DI. So far, a current systematic overview of how often and to what extent teachers use DI for high-achieving students in mixed-ability classrooms is missing (cf. Bondie et al., 2019).

Research has shown that *practitioners' beliefs and expectations* (e.g., importance or usefulness of instructional practices) were related to the (intended) use of different instructional practices in the classroom (Bondie et al., 2019; Gebauer & McElvany, 2017; Trivette et al., 2012). The empirical evidence for instructional practices alone does not automatically lead to their implementation into practice (Missett et al., 2014). Therefore, besides the impact of DI on high-achieving students' outcomes, how teachers and high-achieving students evaluate DI's utility for promoting high-achieving students in mixed-ability classrooms is important, as it could influence the probability of its implementation. As students and teachers might focus on different aspects of DI, both perspectives are valuable and complement each other (Kunter & Baumert, 2006). Studies from the 1990s already hinted that teachers and high-achieving students perceived DI as a useful instructional approach in mixed-ability classrooms (Friedman & Lee, 1996; Hughes, 1999).

Finally, previous studies regarding DI have shown that teachers need assistance in handling heterogeneity in the classroom. In this context, possible *barriers* occurring during the implementation of DI, for example, time constraints, should be investigated (Bondie et al., 2019; Deunk et al., 2018; Subban, 2006). Teachers who do not feel competent to differentiate instruction will not implement DI in a meaningful way (Dixon et al., 2014). However, previous research also found some *facilitators* for implementing DI, such as colleagues, mentors, and control in decision-making (Bondie et al., 2019; Puzio et al., 2015). Hence, investigating barriers and facilitators while implementing DI for high-achieving students in mixed-ability classrooms might help support successful implementation in the future (Russell, 2018).

Present Study

Despite growing efforts regarding DI, knowledge gaps (e.g., an overview of the impact on motivational-affective characteristics and teachers' subjective experiences with DI) still exist, and the differential impact on high-achieving students is still inconclusive (Deunk et al., 2018; Smale-Jacobse et al., 2019; Subban, 2006).

Instructional approaches, such as DI, may positively affect the whole investigated sample but not necessarily high-achieving students (Connor et al., 2011; Cronbach & Snow, 1977). In addition to DI's differential impact, the extent of use, the perceived utility, and barriers and facilitators for implementing DI have not yet been systematically summarized, focusing on high-achieving students. Thus, in the present study, we conducted a systematic mixed-methods review to address the following four research questions (RQs):

1. Which impact does DI have on high-achieving students regarding their academic achievement and motivational-affective characteristics?
2. To what extent do teachers use DI for high-achieving students?
3. How useful is DI perceived to be for high-achieving students by
 - a. teachers and
 - b. high-achieving students themselves?

4. What are the
 - a. barriers and
 - b. facilitators for school staff in implementing DI for high-achieving students?

To answer these questions, we conducted a systematic mixed-methods review, as it has the potential to answer them more comprehensively, more concretely, and in greater detail by using the advantages of both the qualitative and the quantitative methods (Heyvaert et al., 2013; Smale-Jacobse et al., 2019). High-achieving students are defined as students showing outstanding achievement (National Association for Gifted Children, 2010), which can be operationalized in different ways. We used a broad definition of high-achieving students to include students who showed high academic achievement relative to their classmates (e.g., top third or top quarter of the class achievement distribution) and gifted students (e.g., defined through an IQ > 130). This definition was appropriate in our study since our focus was on high-achieving students in mixed-ability classrooms rather than gifted students in specialized gifted programs. We excluded studies investigating *only* underachieving students as our systematic review focused on mixed-ability classrooms and because underachieving students differ in their characteristics from high-achieving students (Agalotis & Kalyva, 2019; McCoach & Siegle, 2003). Finally, we investigated primary and secondary education since DI is important in both settings. In untracked primary education, students' heterogeneity is inherently high, and also in tracked secondary education, a large heterogeneity regarding academic achievement still exists (Bohl et al., 2012; Deunk et al., 2018; Smale-Jacobse et al., 2019).

Method

Study Search

To find relevant literature, we conducted a two-step search. First, we searched journal articles in three educational databases named ERIC, Web of Science Core Collection, and PsycINFO. ERIC contains topics from the psychological and pedagogical research field, focusing on education topics, whereas Web of Science Core Collection is an interdisciplinary database. PsycINFO focuses on psychology and pedagogy.

As our scope was to find any DI approach used in mixed-ability classrooms to promote high-achieving students, we used broad search terms describing the population (e.g., gifted), the intervention (e.g., differentiat*), and student outcomes (e.g., achievement; cf. Petticrew & Roberts, 2006). The full search syntax can be found in Supplement 1.

Besides the first database search, we conducted a so-called *snowball search* and screened the included journal articles' lists of references from the first database search for more relevant studies. Additionally, a hand search was conducted in three established journals: *Gifted Child Quarterly*, *Journal for the Education of the Gifted*, and *Journal of Advanced Academics*. We limited the hand search for pragmatic reasons to journal articles published between January 2013 and September 2019 (cf. Singer & Alexander, 2017).

In a second step, we expanded our first search and conducted a second database search with three different database-specific search-word combinations acquired from each database's thesaurus. As Web of Science does not employ a thesaurus, we chose, for the second search, the databases ERIC, PsycINFO, and Education Source. The first part of the syntax described the population (e.g., gifted children), and the second part the intervention (e.g., individualized instruction; cf. Supplement 1 for details). We intentionally did not define any outcomes in this search, as we aimed to include academic achievement and motivational-affective outcomes as both may play an important role in the education of high-achieving students. These searched databases included, next to journal articles, research reports, conference contributions, and theses, which were all screened and coded.

Inclusion and Exclusion Criteria

All studies included in the present systematic review had to meet the following criteria:

1. Studies were related to DI for high-achieving students inside mixed-ability classrooms. Studies *only* related to structural differentiation (e.g., tracking), out-of-school programs (e.g., summer programs), or other special settings *only* used for high-achieving students (e.g., high-achieving classes) were excluded.
2. Studies provided qualitative and/or quantitative primary data to at least one of the four investigated RQs. Research syntheses and nonempirical work were excluded.
3. The studies' samples were school staff (i.e., teachers, principals, and school coordinators responsible for implementing intervention programs) and/or high-achieving students *without* further special educational needs, so-called twice-exceptionals. The study's focus was *not only* on underachieving students.
4. Studies were related to primary education (grades 1–4) or secondary education (grades 5–13) and *not* to higher education or preschool education only.
5. We included studies taking place in all major academic subjects of schooling (e.g., mathematics, science, reading) and excluded studies focusing merely on religious, physical, art, or another sort of artistic instruction.
6. Studies were published between 2000 and September 2019. In order to summarize the most current research regarding this topic that is not outdated for the current mixed-ability classroom situation, we decided to include studies from 2000 onward (see also Barbier et al., 2022; Neuendorf et al., 2022).
7. Studies' full texts were available in English.

Study Screening and Coding

We conducted two stages of screening. In the first stage, all references were screened by title and abstract according to the above-mentioned inclusion and exclusion criteria. The remaining references were included in the second stage of screening, in which the

full texts were searched and afterward screened for eligibility. Further, a detailed coding form (Excel spreadsheet) was developed for the coding step and piloted with studies excluded from the final dataset. The coding form included the following major characteristics of *all* eligible studies: country, RQs, research approach (i.e., qualitative, quantitative, or mixed-methods), and methods (e.g., interview, achievement test), sample characteristics (e.g., size, grade level, identification method), level of the education system (i.e., primary or secondary), DI type (i.e., grouping, tiered activities, part of a broader program, other, DI in general), and school subject. Studies investigating DI's impact on high-achieving students' outcomes were coded with DI type *others* when their investigated DI type did not fit into one of the three categories, namely tiered activities, grouping, or DI as part of a broader program. Studies that investigated DI independent of a certain instructional approach but focused on DI, in general, were coded with the DI type *DI in general*. For studies investigating the impact of DI on students' outcomes, we additionally coded outcome measures, control group intervention, study duration, and occurrence of teacher training. Missing information needed for effect size calculation was requested via email from the studies' corresponding authors. Missing information regarding all other variables was coded as *missing (m)*.

We further conducted a study quality assessment for all studies contributing to RQ 1 to rate the quality of evidence investigating DI's impact on high-achieving student outcomes. As a first step, we included all studies independent of their study quality and design in the review, as an exclusion based only on these criteria can cause bias in a systematic review (Petticrew & Roberts, 2006). Based on this study quality assessment, we aimed to describe the quality of the evidence base to assess how trustworthy the results are. We further reported only studies in the results section that met more than half of the applicable study quality criteria. The study quality assessment also helped to suggest which further research in this field is needed (Booth et al., 2016; Tod et al., 2022). The code *partially* was assigned when some information regarding the criterion was provided, but the criterion was not fully met. If, for example, the information is given that high-achieving students are identified through a standardized achievement test, but the concrete test and the cutoff score is not mentioned, the code *partially* was assigned. Otherwise, *yes* was assigned when the criterion was fully met and *no* when the criterion was not met.

Interrater Agreement

The studies were screened (eligibility) and coded (information extraction) by three research associates and two research assistants. We double-coded approximately 20% of the references in the screening step. In our study, the inclusion versus exclusion screening was shifted towards exclusion, wherefore Yules Y is a useful statistic to calculate the interrater agreement. Yules Y was $Y = 0.83$ in our study and can be rated as a very good interrater agreement (Wirtz & Kutschmann, 2007).

We double-coded approximately 57% of the included studies in the coding step. Cohen's Kappa was calculated separately for each of the seven closed-coded variables.

Cohen's Kappa ranged between $\kappa = 0.67$ (identification method) and $\kappa = 0.91$ (research approach), indicating a substantial to a nearly perfect interrater agreement for this coding step (Landis & Koch, 1977). During the screening and coding process, regular meetings were held to avoid coder drift, to discuss open questions, and resolve all disagreements through discussion.

Study Analysis

In our systematic review, DI's impact is defined as a positive, neutral, or negative effect (effectiveness) in quantitative studies. A positive effect can, for example, represent a larger pretest–posttest gain in student outcomes for students learning with DI compared to students learning without DI (pretest–posttest-control design) or a positive gain from pretest to posttest (pretest–posttest design). In qualitative studies, we focused on the results and interpretations of the primary studies' authors and defined impact through a positive reported outcome as, for example, *good work* or *improved achievement*. Overall, the numerical and verbal results provide a comprehensive picture of the impact of DI on high-achieving students.

Effect sizes have been computed from studies with a pretest–posttest-control design (RQ 1). Therefore, d_{ppc2} with the difference of mean change for treatment and control group, the pooled pretest standard deviation, and a correction for small sample sizes was calculated (Morris, 2008). The effect size and sampling variance calculations were conducted using the software R and the package metafor (Viechtbauer, 2010, 2021). As the number of studies using pretest–posttest-control design was small ($n = 6$ with 30 effect sizes) and the investigated DI types, as well as outcome measures, were quite heterogeneous, we did not conduct a quantitative meta-analysis, as it would not be appropriate with this heterogeneous set of studies (cf. Deeks et al., 2019; Noetel et al., 2019; Siddaway et al., 2019).

From studies investigating the impact of DI on student outcomes (RQ 1) with no pretest–posttest-control group design, we extracted the main findings through a qualitative synthesis approach. The first step included “intensive, repetitive, and above all highly active reading” to familiarize with the studies and extract relevant quotes (Lee et al., 2015, p. 341). Subsequently, these quotes were sorted and narratively summarized.

For RQs 2–4, a thematic synthesis was conducted. This approach can investigate individuals' perspectives and experiences, for example, regarding barriers and facilitators. The approach aims to identify different influencing factors of a certain topic and organize those into common themes (Booth et al., 2016). Therefore, we actively read all relevant primary studies (quantitative, qualitative, and mixed-methods) and extracted the main findings through quotes. Afterward, we sorted and summarized them. Commonalities and contrasts between the primary studies were detected. Thereof, we created themes through an inductive approach without defining them beforehand (cf. Lindner & Schwab, 2020). Further, according to Petticrew and Roberts (2006), an overview with extensive information on the studies was created. Different quotes from primary studies illustrate and clarify the results and give

information about the study contexts. For RQ 4, a graphical representation displaying the barriers and facilitators was created.

Results

Descriptive Overview of the Included Studies

Through both database searches and the hand search, 7,680 references were detected and screened regarding their title and abstract. From these references, 785 references were included in the full-text screening. The hand search revealed three studies, and the two database searches 44 studies that were finally included in our systematic review. An additional 56 references were detected through the snowball search. From these 56 references included in the full-text screening, two studies were finally included in our systematic review. Overall, 49 references were included and analyzed in the current systematic review (cf. Figure 1). Detailed information about the study characteristics (e.g., publication year, DI type, identification method) can be found in Table 1 and Tables S1–S5.

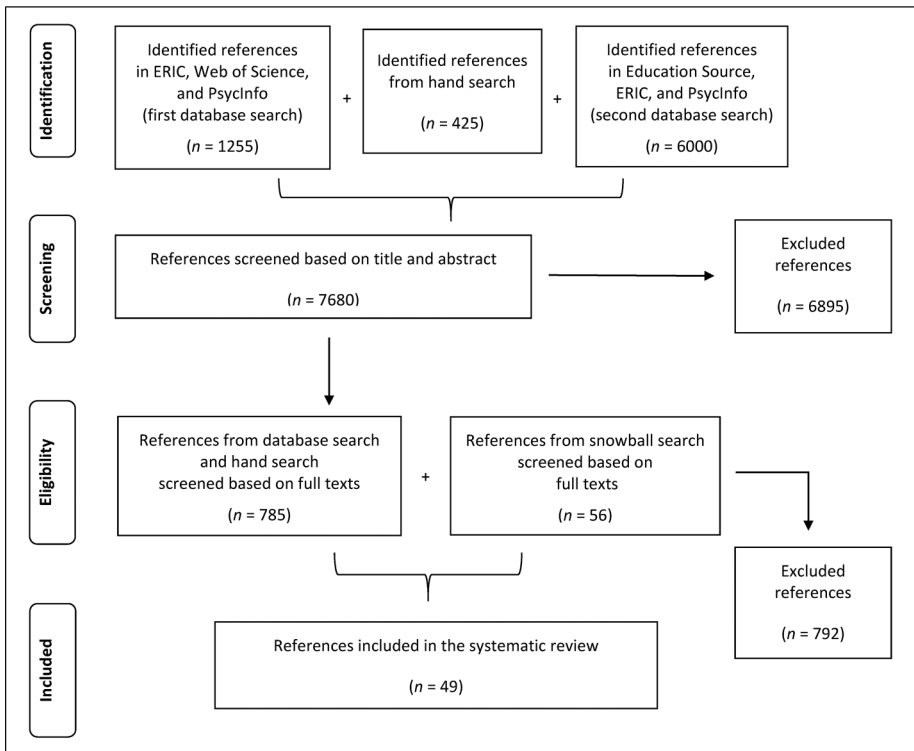


Figure 1. Flow diagram describing the study search and selection process.

Impact of DI on High-Achieving Student Outcomes (RQ 1)

Overall, 15 studies (Table S1) investigated the impact of DI on students' academic achievement and/or motivational-affective characteristics. First, the study quality assessment results (Table S2) are reported to get an overview of the quality of the evidence. Second, studies investigating the impact of DI on high-achieving students' academic achievement and motivational-affective characteristics are summarized (Table S1). Only studies that met more than half of the applicable study quality criteria were included in the second section.

Quality of evidence for the impact of DI on high-achieving students' outcomes

The quality assessment of the primary studies included for RQ 1 is displayed in Table S2. Overall, no study met all investigated study quality criteria. However, one study came very close (Shaunessy-Dedrick et al., 2015). Three studies (Kim et al., 2014; Nomi, 2010; Ysseldyke & Tardrew, 2007) were not considered for further presentation of results because they did not meet more than half of the applicable study quality criteria. Overall, out of the remaining 12 studies, all studies, except one, reported a clear operationalization of DI. Further, except one, all studies reported the exact sample size of the high-achieving students. Four studies did not report a high-achieving student comparison group without DI. From eight studies with high-achieving student comparison groups, six studies did not investigate the comparability between the two groups *before* DI was used. Further, a randomized allocation to the experimental and comparison group at the student level was not conducted in six out of eight studies with comparison groups.

Evidence regarding the impact of DI on high-achieving students' outcomes

As shown in Table 2, six pretest–posttest–control design studies (Delcourt et al., 2007; Meyer et al., 2011; Robinson et al., 2014; Saleh et al., 2005; Shaunessy-Dedrick et al., 2015; Ysseldyke et al., 2004) investigated the impact of DI on high-achieving student outcomes compared with high-achieving students without DI. Overall, those studies included 30 effect sizes, from which 20 effect sizes represented academic achievement differences. A positive effect size represents an advantage of the pretest–posttest gain for the experimental group with DI compared to the control group without DI. In contrast, a negative effect size represents an advantage for the control group. The effect sizes ranged for academic achievement from $d_{ppc2} = -0.692$ to $d_{ppc2} = 3.155$ (without one outlier: $d_{ppc2} = 7.179$). Although a mixed picture of academic achievement outcomes appeared, more positive ($n = 13$) than negative effects ($n = 7$) were found. Ten effect sizes represented motivational-affective differences. These ranged from $d_{ppc2} = -0.562$ to $d_{ppc2} = 0.736$. Positive effects ($n = 6$) outweighed negative effects ($n = 4$).

Further, three posttest only with control design studies¹ (McCoach et al., 2014; Meyer et al., 2011; Tieso, 2005), two qualitative or mixed-methods studies (Bellamy, 2005; Dreeszen, 2009), and two studies with other research designs

Table 1. Descriptive Statistics of the Included Primary Studies (*n* = 49).

		Number of studies
Publication year	2000–2009	18
	2010–2019	31
Country	USA or Canada	30
	Great Britain & Ireland	4
	Netherlands	3
	Australia	2
	Germany	2
	Turkey	2
	Other	6
Type of DI	DI in general	27
	DI as part of a broader program	9
	Grouping	5
	Tiered activities	4
	Other	4
Level of the education system	Primary education (grades 1–4)	14
	Secondary education (grades 5–13)	10
	Primary and secondary education	24
	Information missing	1
Identification method for high-achieving students (only studies including a student sample, <i>n</i> = 25)	Multiple methods (e.g., standardized tests, creativity measures, teacher recommendations)	10
	Standardized tests	11
	Teacher recommendations	2
	Information missing	2
Student outcomes (only studies investigating DI's impact, <i>n</i> = 15 ^a)	Mathematics	6
	Science	4
	Reading	4
	Motivational-affective outcomes	6
	Other	3
Teacher training regarding DI (only studies investigating DI's impact, <i>n</i> = 15)	Yes	10
	No	5

^aAs most studies investigated various outcomes, the sum does not add up to 15.

Table 2. Pretest–Posttest–Control Group Studies Investigating DI.

Author, year	Experimental group intervention	Control group intervention	Identification criteria—high-achieving students	Student outcome	Effect size d_{ppz2}	Sampling variance ^a
Delcourt et al., 2007	Other: different DI practices (e.g., independent study)	High-achieving students without a program for the gifted at their school	Multiple methods: EG: IQ, achievement, teacher evaluation, parent evaluation, and student evaluation KG: teacher nomination, largely based on reading and mathematics achievement	Mathematics concepts	−0.347	0.026
				Mathematics problem solving	−0.093	0.026
Meyer et al., 2011	Tiered activities: web-based training with individualization, including remediation and enrichment	Web-based training without individualization	Standardized achievement test: upper third of the standardized pretest scores on the reading comprehension test	Reading comprehension	−0.305	0.026
				Science	−0.239	0.026
				Social studies	−0.692	0.027
				Scholastic competence	0.276	0.029
				Social acceptance	−0.562	0.030
				Internal motivation	0.293	0.028
				Independent judgment	0.294	0.030
				Independent mastery	0.736	0.030
				Preference of challenge	0.198	0.029
				Reading comprehension 1: standardized test	0.125	0.070
Reading comprehension 2: total recall	0.134	0.073				

(continued)

Table 2. (continued).

Author, year	Experimental group intervention	Control group intervention	Identification criteria—high-achieving students	Student outcome	Effect size d_{ppc2}	Sampling variance ^a
				from comparison text		
				Reading comprehension 3: use of structure strategy in comparison text	0.136	0.073
				Reading comprehension 4: total recall from problem and solution text	-0.161	0.073
				Reading comprehension 5: use of structure strategy in problem and solution text	-0.194	0.073
				Reading comprehension 6: fill in missing signaling words	0.531	0.076

(continued)

Table 2. (continued).

Author, year	Experimental group intervention	Control group intervention	Identification criteria—high-achieving students	Student outcome	Effect size d_{ppz2}	Sampling variance ^a
Robinson et al., 2014	Part of a broader program: STEM Starters intervention, a differentiated, problem-based science curriculum	Science instruction according to the school-adopted science curriculum	Multiple methods: standardized achievement test scores, cognitive ability tests (verbal and nonverbal), creativity measures, teacher recommendations, parent recommendations	Science process skills year 1 Science content knowledge year 1 Science concept knowledge year 1	1.127 3.155 2.089	0.058 0.121 0.077
Saleh et al., 2005	Grouping: Homogeneous grouping	Heterogeneous grouping	Standardized achievement Test: upper 25% on Science Elementary Achievement Test (SEAT)	Science process skills year 2 Science content knowledge year 2 Science concept knowledge year 2 Science achievement: biology Motivational beliefs 1: academic benefits Motivational beliefs 2: social benefits Motivational beliefs 3: attitude benefits	0.426 7.179 0.856 0.725	0.029 0.353 0.044 0.121

(continued)

Table 2. (continued).

Author, year	Experimental group intervention	Control group intervention	Identification criteria—high-achieving students	Student outcome	Effect size d_{ppc2}	Sampling variance ^a
Shaunnessy-Dedrick et al., 2015	Part of a broader program: Schoolwide Enrichment Model—Reading (SEM-R), a differentiation approach divided into three phases	District's typical reading instruction	Multiple methods: IQ, group screening, teacher checklist, alternative identification criteria	Reading comprehension <i>Attitudes toward reading</i>	0.194 -0.068	0.141 0.135
Ysseldyke et al., 2004	Part of a broader program: Accelerated Math, a curriculum-based instructional management system	Regular math instruction	m^b	Mathematics achievement	0.411	0.033

Note. Motivational-affective student outcomes are presented in italics.

^aSampling variances were calculated according to Viechtbauer (2021). Because only one study gave the correlation between pre- and posttest (Shaunnessy-Dedrick et al., 2015), we used $r = .60$ as a conservative estimate for the studies that did not report the correlation (cf. van Alren et al., 2019 for a similar procedure). ^b m stands for missing information, meaning that the concrete identification criteria for the high-achieving sample have not been reported.

(Faber et al., 2018; Kondor, 2007) investigated the impact of DI on high-achieving student outcomes (Table S1). As pretest differences respectively, control group outcomes cannot be taken into account with all of these designs, the impact should be interpreted with some caution. Altogether, these studies reinforced the positive impact of DI on high-achieving students' academic achievement and motivational-affective characteristics in mixed-ability classrooms. In an exemplary mixed-methods study investigating the impact of DI on high-achieving students, it was summarized that "differentiation by task seemed to provide an effective learning opportunity, which in turn produced good work and had apparently stretched the abilities of the students" (Bellamy, 2005, p. 79). Further, high-achieving students taught with DI also "expressed a desire to continue with the differentiated reading program" (Dreeszen, 2009, p. 312).

The Extent of DI Use (RQ 2)

Overall, 12 studies (Table S3) provided information concerning the extent of DI. When teachers were asked how frequently they explicitly consider high-achieving students in their mixed-ability classrooms, many teachers stated that they do not actively consider these students and make no or only minor changes in their regular curriculum to meet high-achieving students' needs (Abu et al., 2017; Al-Lawati & Hunsaker, 2007; Brighton et al., 2007; Endepohls-Ulpe & Thömmes, 2014; Laine & Tirri, 2016; Prast & de van Weijer-Bergsma, 2015).

Classroom observations and audio recordings from classroom conversations in mixed-ability classrooms complemented these self-reported teachers' views. Findings showed that most high-achieving students did not get DI regularly (Gilson et al., 2014; Reis et al., 2004). Reis et al. (2004, p. 325) summarized the key finding of their study as follows: "The major finding and core category in this study was the absence of differentiated instruction for talented readers; talented readers received some challenge in three of the participating classrooms, but limited opportunities in the other nine."

One study asked school representatives (e.g., gifted education coordinators, administrators for curriculum and teaching) about DI frequency. Here, a somewhat different picture was painted. 86% responded that DI for high-achieving students in the regular classroom exists but also mentioned that the actual use varies between teachers (Jarvis & Henderson, 2015).

Three other studies asked high-achieving students about the frequency of DI in their mixed-ability classroom. In two studies from the United States, more than two-thirds of the students said that their instruction is not differentiated and that they learn the same content, at the same level, and the same pace as the other students (Assouline et al., 2013; Swiatek & Lupkowski-Shoplik, 2003). In the third study, from Turkey, the picture was somewhat reversed when students were asked if they can work with their intellectual peers and if the lessons are according to their needs, interests, and knowledge. Here, approximately 60%–80% of the asked students said they received DI in five different subjects (Kahveci & Akgül, 2014).

To summarize, teachers and high-achieving students mainly reported that DI is not used for high-achieving students in mixed-ability classrooms proactively and on a

regular basis. Nevertheless, in some classes, teachers differentiated instruction according to the high-achieving students' needs. From these findings, the question arises if teachers make these limited modifications because they see low utility in DI or if other barriers hinder the implementation.

Utility of DI (RQ3)

Overall, 13 studies (Table S4) investigated how useful teachers and high-achieving students rated DI. Thereof, 10 studies (Apps, 2011; Brevik et al., 2018; Dreeszen, 2009; Endepohls-Ulpe, 2017; Endepohls-Ulpe & Thömmes, 2014; Oswald & Villiers, 2013; Parish, 2016; Reis et al., 2010; Simmons, 2018; Vreys et al., 2018) provided information on teachers' perspectives, and three studies (Adams-Byers et al., 2004; Dobron, 2011; Kanevsky, 2011) provided information on students' perspectives.

Teacher Perspective

When teachers were asked if DI is a useful approach for high-achieving students' promotion, they mainly agreed and acknowledged the value and need of DI for providing a high-quality curriculum (Apps, 2011; Brevik et al., 2018; Endepohls-Ulpe, 2017; Endepohls-Ulpe & Thömmes, 2014; Oswald & Villiers, 2013; Simmons, 2018). One study summarizes this result by stating that "internal differentiation was clearly rated to be not only the most effective one but also to be the one with hardly any negative consequences" for students (Endepohls-Ulpe, 2017, p. 159). Further, a positive relationship between the frequency of use and the rated eligibility of DI was found. In contrast, a negative relationship was found between frequency of use and estimated workload (Endepohls-Ulpe & Thömmes, 2014). Further, teachers stated that two-thirds preferred homogeneous grouping over heterogeneous grouping for teaching if they had to prioritize between homogeneous and heterogeneous grouping. Teachers' main reason was that homogeneous grouping could promote a more thorough academic development for their high-achieving students (Apps, 2011).

In four other studies, teachers implemented specific DI types and reported how they perceived their utility. In general, teachers stated positive outcomes of DI and were impressed by the positive outcomes (e.g., higher motivation and involvement) they perceived for high-achieving students (Dreeszen, 2009; Parish, 2016; Reis et al., 2010; Vreys et al., 2018). The following quote illustrates which positive outcomes teachers stated: "Children [...] were eager to work on the enrichment activities and were notably less bored. [...] The children showed increased self-confidence, better working attitude, and improved behavior and well-being" (Vreys et al., 2018, pp. 12–13).

Student Perspective

Regarding the high-achieving students' perspective toward DI, it could be summarized that their attitude was positive in general and that they supported its use (Dobron, 2011; Kanevsky, 2011). High-achieving students "perceived the differentiated curriculum as

appropriately challenging, worthy of selection, appealing to interests, satisfying as a learning experience, useful in study, and applicable to the real world.” (Dobron, 2011, p. 248). In one study, students were asked about the advantages of homogeneous vs. heterogeneous grouping. They saw advantages of mixed-ability grouping regarding social and emotional aspects and advantages of high-ability grouping regarding academic aspects (Adams-Byers et al., 2004). This perspective regarding grouping is in accordance with the teachers’ perspective reported above.

Barriers and Facilitators for Implementing DI (RQ 4)

Overall, 22 studies (Table S5) contributed to answering the question of barriers and facilitators for implementing DI for high-achieving students in mixed-ability classrooms. Figure 2 gives an overview of the detected barriers (left) and facilitators (right).

Barriers

Seven categories of barriers repeatedly occurred in the investigated studies. First, a lack of *time* for preparing and implementing DI was reported by the school staff (e.g., Abu et al., 2017; Brighton et al., 2007; Coates, 2009; Cross et al., 2018; Endepohls-Ulpe, 2017; Johnsen et al., 2002; Reis et al., 2004; Rubenstein et al., 2015; Russell, 2018; Seedorf, 2014; Vreys et al., 2018). Second, a lack of *resources*, including, among others, material and financial resources, was reported (e.g., Abu et al., 2017; Apps, 2011; Cross et al., 2018; Johnsen et al., 2002; Oswald & Villiers, 2013; Reis et al., 2004). Third, a lack of *knowledge and training* was repeatedly mentioned, as teachers did not feel sufficiently informed (e.g., Brevik et al., 2018; Brighton et al., 2007; Cross et al., 2018; Jarvis & Henderson, 2015; Oswald & Villiers, 2013; Reis et al., 2004; Rubenstein et al., 2015; Seedorf, 2014). Fourth, teachers reported *misconceptions* regarding high-achieving students and DI

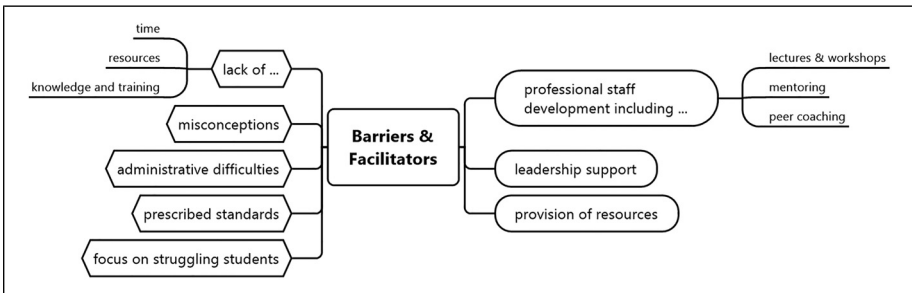


Figure 2. Barriers and facilitators for implementing differentiated instruction for high-achieving students in mixed-ability classrooms. Note. The hexagons on the left side represent barriers, and the ovals on the right side represent facilitators for implementing differentiated instruction.

(e.g., Abu et al., 2017; Brighton, 2003). Abu et al. (2017, p. 101), for example, summarized that “A great majority of the teachers in study group think that changes are not made during the teaching program because [...] gifted children are already good, small in number and do not need a different education.”

Fifth, school staff reported *administrative difficulties*, including organizational aspects of schools and the classroom, such as overcrowded classrooms and management issues, as well as challenges in reliably identifying high-achieving students and assigning them to adequate DI without misplacement (e.g., Brighton et al., 2007; Dreeszen, 2009; Endepohls-Ulpe, 2017; Johnsen et al., 2002; MacIntyre & Ireson, 2002; Oswald & Villiers, 2013; Rubenstein et al., 2015). The sixth category includes barriers through prescribed (state) *standards*. Some school staff reported less flexibility in differentiating instruction (e.g., Abu et al., 2017; Reis et al., 2004; Rubenstein et al., 2015). The following quote, from a study that investigated how teachers react to DI, emphasizes this barrier: “Almost every teacher [...] expressed either frustration with the overemphasis of standards or the lack of a connection between state and study expectations” (Rubenstein et al., 2015, p. 157). The last category contains the barrier that the focus for support measures, including DI, is on *struggling students* instead of high-achieving students (e.g., Apps, 2011; Brighton et al., 2007; Cross et al., 2018; Reis et al., 2004). The last two categories partly showed overlap insofar that some (state) standards suggested focusing on struggling students instead of focusing on high-achieving students (e.g., Reis et al., 2004).

Facilitators

The most frequently stated facilitator was *professional staff development*, which included, among others, training (e.g., lectures and workshops), mentoring or peer coaching, and provision of material. Further, these professional staff development activities were partly accompanied by *leadership* support and the *provision of resources*. Through these facilitators, teachers were able to gain knowledge, were higher motivated, and felt more confident. In general, these support measures were rated as useful for and facilitated the implementation of DI (e.g., Brighton et al., 2007; Coates, 2009; Johnsen et al., 2002; Latz et al., 2008; Oswald & Villiers, 2013; Prast et al., 2018; Reis et al., 2010; Rubenstein et al., 2015; Seedorf, 2014; Vreys et al., 2018; Yuen et al., 2018). The usefulness of the facilitators mentioned above is illustrated by the following statement, which resulted from a staff development project: “The vast majority of participants viewed staff-development activities, leadership, the mentor teachers, and project support as being beneficial to extremely beneficial” (Johnsen et al., 2002, p. 61). Another study concluded that “many teachers are able to differentiate if provided appropriate support and materials (e.g., preassessments and tiered activities”); Rubenstein et al., 2015, p. 159).

To summarize, various barriers (i.e., lack of time, resources, and knowledge; misconceptions; administrative difficulties; prescribed (state) standards; and the focus on struggling students), making the implementation of DI difficult, were found. Nonetheless, different factors, especially from the area of professional staff

development, were mentioned that supported DI's implementation for high-achieving students in mixed-ability classrooms.

Discussion

The findings of the systematic mixed-methods review are evaluated in the context of other research, and it is discussed how the results relate to each other. Further, limitations, directions for future research, and implications for practice and policy are considered.

Impact of DI on High-Achieving and Not-as-High-Achieving Identified Students

One main finding of the systematic mixed-methods review is the mainly positive impact of DI on high-achieving students in mixed-ability classrooms. These findings are supported through studies using different quantitative designs as well as through studies using qualitative study designs with classroom observations, interviews, and document reviews. Further, the mainly positive impact is in accordance with previous research syntheses, which found, for students in general, mainly small to moderate positive outcomes of DI, including different grouping arrangements, on students' academic achievement (Deunk et al., 2018; Lou et al., 1996; Puzio et al., 2020; Smale-Jacobse et al., 2019). However, the current review found large heterogeneity between studies for comparable outcomes and within studies for different outcomes. This heterogeneity might be, to some part, explainable through the different study designs used and the different study quality criteria met. Delcourt et al. (2007), for example, found negative effects for high-achieving students' achievement outcomes but mainly positive effects for motivational-affective characteristics. As the comparability of the experimental and comparison group was not investigated at the beginning and no randomized allocation to the two groups took place, the baseline level of the two groups at the beginning might differ. Interestingly, no baseline equivalence could be found for the achievement outcomes, in contrast to the motivational-affective characteristics. In cases where baseline equivalence was missing, the experimental group with DI scored higher at the pretest than the control group without DI. Hence, some negative effects might be explained through different baseline levels (pretest scores) in both groups, resulting in ceiling effects of the instruments and/or regression to the mean. Overall, the evidence regarding the impact of DI on high-achieving students in mixed-ability classrooms is sparse and partially of low quality. Therefore, the need for more empirical research with high-quality standards in this field is still high (Barbier et al., 2022; Shaunessy-Dedrick et al., 2015; White et al., 2003).

The Extent of DI Use and Its Perceived Utility

The limited extent of DI for high-achieving students in mixed-ability classrooms is in line with the findings of the Programme for International Student Assessment 2012.

Students were asked if teachers give different work to struggling or high-achieving students in their classroom in mathematics. On average, only 30% of students in the OECD countries declared that it was true in every or most lessons, whereby the values ranged from 13% in France and Italy to 62% in Sweden. Especially from high-achieving students, compared to low-achieving students, student-oriented instruction, including the question mentioned above, was reported less frequently (Schleicher, 2016). One possible explanation for this lack of DI can be that existing schoolbooks included few examples of DI (Heinle et al., 2022). However, teachers reported that they use schoolbooks almost every day (Wenglinsky, 2002), and over 80% reported schoolbooks as the basis for their teaching (Wendt et al., 2017). Thus, teachers need to actively differentiate and adapt their instruction beyond using the given tasks from schoolbooks.

The perceived contradiction that teachers, on average, rate DI for high-achieving students in mixed-ability classrooms as useful but, on average, do not proactively nor regularly implement DI can be explained, for example, as follows: “Teachers know that gifted students should be differentiated, but as they do not receive education and training to address gifted students’ needs, they do not apply the practices known from the literature” (Laine & Tirri, 2016, p. 158). This contradiction is also acknowledged in studies outside of DI. Philippe (2017) investigated the implementation of social and emotional learning and concluded that the knowledge about its utility and importance does not necessarily come along with the confidence to implement it effectively. Moreover, the studies in the current review did not use the same sample to investigate the extent of usage and the perceived utility of DI. One exception is the study by Endepohls-Ulpe and Thömmes (2014), who found a positive correlation between self-reported frequency of use and the perceived utility of DI.

Barriers and Facilitators for Implementing DI

One factor that might influence the perceived utility of DI and its extent of usage is professional development for teachers. Brevik et al. (2018, p. 36) stated that “the literature reviewed on teacher education and classroom practices around the world suggests that developing teachers’ differentiation practice and experience is crucial”. Further, professional development regarding DI was positively associated with higher teacher self-efficacy and efficacy beliefs, and the teacher self-efficacy and efficacy beliefs were positively associated with the implementation of DI (Dixon et al., 2014). On the other hand, a survey conducted by the National Association for Gifted Children and the Council of State Directors of Programs for the Gifted asked if university coursework in gifted education is mandatory for all preservice teachers in the United States. Out of 48 states answering this question, only three states answered *yes*, whereas 45 answered *no* (Rinn et al., 2020). This finding is in accordance with our review that found a perceived lack of training and knowledge regarding DI for teachers.

Overall, the barriers and facilitators found in the current review affirm Sisk (2009), who stated that classroom teachers need support for educating the high-achieving students in mixed-ability classrooms. The regular classroom teachers

can hardly do it on their own without professional development, willingness, a supportive belief system, and school culture (Sisk, 2009). Furthermore, the detected barriers and facilitators are similar to the barriers and facilitators mentioned in the context of DI and students in general, as well as in inclusive settings (Bondie et al., 2019; Lindner & Schwab, 2020; Puzio et al., 2015). Therefore, the barriers and facilitators might be mainly generalizable for implementing DI, independent of the underlying student population. Teachers should be trained to know the individual needs of all students and be supported accordingly with, for example, appropriate supplemental resources.

Limitations and Directions for Future Research

Overall, only a few empirical studies with a suitable research design evaluated the impact of DI on high-achieving student outcomes. This follows the results of a systematic review, which aimed to give an overview of empirical implications for the education of high-achieving students and concluded that there are relatively few studies. Thus, “evidence-based policy and practice are scarce” (White et al., 2003: vii). Another difficulty we encountered is that the included studies differed greatly in study design, operationalization of DI, identification of high-achieving students, and partially included very small samples (Bellamy, 2005; Kondor, 2007). The fact that many studies did not provide detailed information regarding the definition and operationalized features of DI was already stated by a previous synthesis (Deunk et al., 2018). Although we tried to code specific DI features, it was impossible to figure them out due to a lack of studies and information. Accordingly, for future studies, it is very important that primary study authors explicitly state how they operationalized DI—although DI might be a part of a broader program—in order to investigate the different forms, features, and quality of DI (cf. Laine & Tirri, 2016; Smale-Jacobse et al., 2019). In addition, the student outcomes investigated also differed remarkably between studies. This, among other reasons, prevented us from conducting a quantitative meta-analysis. However, we gained hints that DI might be useful for high-achieving students regarding a broad range of outcomes, including academic achievement and motivational-affective outcomes. So far, the consideration of motivational-affective outcomes regarding the impact of DI was mentioned as a research gap (Smale-Jacobse et al., 2019). In future research syntheses with more primary studies, quantitative syntheses for each student outcome separately might be of interest.

In addition, some primary studies did not report how high-achieving students were exactly identified. Accordingly, in future studies, the authors should report in detail how high-achieving students are exactly identified and their respective sample sizes. As in the educational context, randomization at the student level is often not possible, it is even more important that pretest data for all investigated student subgroups is described and the comparability between the experimental and comparison group is tested beforehand to ensure baseline equivalence. Further, when students are nested into classes, the dependency of their data should be taken into account for statistical

analyses (Hedges, 2007; Hedges et al., 2010). In future research syntheses, with more (comparable) studies investigating DI's impact on high-achieving students' outcomes, the statistical investigation of the effects of various DI implementation conditions, differences in the operationalization of high-achieving students, and the primary studies' quality on DI's impact on high-achieving students outcomes might be informative. A detailed investigation of contextual factors (e.g., socioeconomic status and achievement at the school level) should also be of interest for future research, as existing research provided evidence that contextual factors can influence the effects of DI on student outcomes (McCoach et al., 2014).

Conclusion and Implications for Practice and Policy

The findings of this review are important for researchers, teachers, and policymakers, in influencing their decisions regarding the promotion of high-achieving students. Based on theoretical concepts (Kaplan, 1986; Vygotsky, 1978), previous research (Barbier et al., 2022; Bondie et al., 2019; Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019), and the current systematic mixed-methods review, it can be summarized that matching learning with students' ability levels through DI is a promising approach for the promotion of high-achieving students in mixed-ability classrooms. The criticism that special education for high-achieving students is elitist and unfair does not hold for DI, as DI also positively impacted not-as-high-achieving identified students (Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019). Nonetheless, the implementation of DI alone is not a sure-fire success. This is also in accordance with the claims of Dumont (2018), who stated that lasting changes at all levels of education systems are needed for the implementation of individualized promotion, and with the results of Deunk et al. (2018), who found that DI in the context of a broader educational program is more effective than homogeneous ability grouping alone. We showed that facilitators, such as leadership support, professional staff development, and resource provision (e.g., time, materials), are important for a successful implementation. To achieve this, embedding DI into a broader educational program can be helpful so that changes in different areas, such as staff development, provision of resources, and adaptation of the school mission statement, are made possible and therefore facilitate the implementation of DI.

Acknowledgements

The authors thank Sarah Reinhold, a research associate at the Centre for International Student Assessment (ZIB), and our student assistants for the support in coding our studies.


Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. This work was supported by the Federal Ministry of Education and Research (BMBF), Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK) (grant number ZIB2022).

ORCID iD

Lisa Ziernwald  <https://orcid.org/0000-0002-7299-6367>

Supplemental Material

Supplemental material for this article is available online.

Notes

- 1 Meyer et al. (2011) is also included in the section *pretest–posttest-control group design*, as different designs for the outcomes were used.

References

References marked with an asterisk were included in the systematic mixed-methods review.

- *Abu, N. K., Akkanat, Ç., & Gökdere, M. (2017). Teachers' views about the education of gifted students in regular classrooms. *Turkish Journal of Giftedness and Education*, 7(2), 87–109.
- *Adams-Byers, J., Whitsell, S. S., & Moon, S. M. (2004). Gifted students' perceptions of the academic and social/emotional effects of homogeneous and heterogeneous grouping. *Gifted Child Quarterly*, 48(1), 7–20. <https://doi.org/10.1177/001698620404800102>
- Agaliotis, I., & Kalyva, E. (2019). Motivational differences of Greek gifted and non-gifted high-achieving and gifted under-achieving students. *International Education Studies*, 12(2), 45–56. <https://doi.org/10.5539/ies.v12n2p45>
- *Al-Lawati, F. A. K., & Hunsaker, S. L. (2007). Differentiation for the gifted in American Islamic schools. *Journal for the Education of the Gifted*, 30(4), 500–518. <https://doi.org/10.4219/jeg-2007-501>
- *Apps, K. A. (2011). *Relationships between teachers' perceptions of gifted program status and instructional choices* [Doctoral dissertation, University of Southern California]. ProQuest. <https://www.proquest.com/openview/47b5550fa75c7cc4b7f28d76742436c8/1?pq-origsite=gscholar&cbl=18750>
- Archambault, F. X., Westberg, K. L., Brown, S. W., Hallmark, B. W., Zhang, W., & Emmons, C. L. (1993). Classroom practices used with gifted third and fourth grade students. *Journal for the Education of the Gifted*, 16(2), 103–119. <https://doi.org/10.1177/016235329301600203>
- *Assouline, S. G., Colangelo, N., Heo, N., & Dockery, L. (2013). High-ability students' participation in specialized instructional delivery models: Variations by attitude, grade, gender, and content area. *Gifted Child Quarterly*, 57(2), 135–147. <https://doi.org/10.1177/0016986213479654>
- Bal, A. P. (2016). The effect of the differentiated teaching approach in the algebraic learning field on students' academic achievements. *Eurasian Journal of Educational Research*, 16(63), 185–204. <https://doi.org/10.14689/ejer.2016.63.11>

- Barbier, K., Struyf, E., Verschueren, K., & Donche, V. (2022). Fostering cognitive and affective-motivational learning outcomes for high-ability students in mixed-ability elementary classrooms: A systematic review. *European Journal of Psychology of Education*. Advance Online Publication. <https://doi.org/10.1007/s10212-022-00606-z>
- *Bellamy, L. (2005). A critical analysis of how differentiation can promote the full inclusion of three gifted and talented students in a mixed-ability, year 9 class studying Macbeth. *English Teaching: Practice and Critique*, 4(2), 72–83.
- Betts, G. (2004). Fostering autonomous learners through levels of differentiation. *Roeper Review*, 26(4), 190–191. <https://doi.org/10.1080/02783190409554269>
- Bohl, T., Batzel, A., & Richey, P. (2012). Öffnung - Differenzierung - Individualisierung - Adaptivität. Charakteristika, didaktische Implikationen und Forschungsbefunde verwandter Unterrichtskonzepte zum Umgang mit Heterogenität [Opening - differentiation - individualization - adaptivity. Characteristics, didactical implications and research findings of related teaching concepts for dealing with heterogeneity]. In T. Bohl, M. Bönsch, M. Trautmann, & B. Wischer (Eds.), *Binnendifferenzierung: Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht* (pp. 40–69). Prolog.
- Bondie, R. S., Dahnke, C., & Zusho, A. (2019). How does changing “one-size-fits-all” to differentiated instruction affect teaching? *Review of Research in Education*, 43(1), 336–362. <https://doi.org/10.3102/0091732X18821130>
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Sage.
- Borland, J. H. (2005). Gifted education without gifted children: The case for no conception of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 1–19). Cambridge University Press.
- *Brevik, L. M., Gunnulfson, A. E., & Renzulli, J. S. (2018). Student teachers’ practice and experience with differentiated instruction for students with higher learning potential. *Teaching and Teacher Education*, 71, 34–45. <https://doi.org/10.1016/j.tate.2017.12.003>
- *Brighton, C. M. (2003). The effects of middle school teachers’ beliefs on classroom practices. *Journal for the Education of the Gifted*, 27(2–3), 177–206. <https://doi.org/10.1177/016235320302700205>
- *Brighton, C. M., Moon, T. R., Jarvis, J. M., & Hockett, J. A. (2007). *Primary grade teachers’ conceptions of giftedness and talent: A case-based investigation*. The National Research Center on the Gifted and Talented. <https://www.semanticscholar.org/paper/Primary-Grade-Teachers%27-Conceptions-of-Giftedness-A-Brighton-Moon/7c642a37586670596ac4eade97aaacec90bfda7a>
- *Coates, D. (2009). Developing challenging science activities for gifted pupils through action research. *Education 3–13*, 37(3), 259–268. <https://doi.org/10.1080/03004270802596915>
- Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P. S., Bayraktar, A., Crowe, E. C., & Schatschneider, C. (2011). Testing the impact of child characteristics × instruction interactions on third graders’ reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, 46(3), 189–221.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Irvington Publishers.
- *Cross, T. L., Cross, J. R., & O’Reilly, C. (2018). Attitudes about gifted education among Irish educators. *High Ability Studies*, 29(2), 169–189. <https://doi.org/10.1080/13598139.2018.1518775>
- Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, & V. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 241–284). The Cochrane Collaboration; Wiley Blackwell.

- *Delcourt, M. A. B., Cornell, D. G., & Goldberg, M. D. (2007). Cognitive and affective learning outcomes of gifted elementary school students. *Gifted Child Quarterly*, 51(4), 359–381. <https://doi.org/10.1177/0016986207306320>
- Deunk, M. I., Smale-Jacobse, A. E., Boer, H. d., 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>
- Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, 37(2), 111–127. <https://doi.org/10.1177/0162353214529042>
- *Dobron, K. H. (2011). *Gifted students' perceptions of the differentiated curriculum* [Doctoral dissertation, University of Southern California]. ProQuest. <https://www.proquest.com/openview/70ce36f27fe13b80f6b46c674bf959f7/1?pq-origsite=gscholar&cbl=18750>
- *Dreeszen, J. L. (2009). *The impact of differentiation on the critical thinking of gifted readers and the evolving perspective on the fifth grade classroom teacher* [Doctoral dissertation, Kansas State University]. ProQuest. <https://www.proquest.com/docview/304911139>
- Dumont, H. (2018). Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht [Adaptive teaching: conceptual reflections]. *Zeitschrift für Erziehungswissenschaft*, 22(2), 249–277. <https://doi.org/10.1007/s11618-018-0840-0>
- *Endepohls-Ulpe, M. (2017). Acceleration, enrichment, or internal differentiation—consequences of measures to promote gifted students anticipated by German secondary school teachers. *Electronic Journal of Research in Educational Psychology*, 15(1), 147–163. <https://doi.org/10.14204/ejrep.41.1517>
- *Endepohls-Ulpe, M., & Thömmes, N. (2014). Chances and limitations of implementing measures of differentiation for gifted children in primary schools: The teachers' part. *Turkish Journal of Giftedness and Education*, 4(1), 24–36.
- *Faber, J. M., Glas, C. A. W., & Visscher, A. J. (2018). Differentiated instruction in a data-based decision-making context. *School Effectiveness and School Improvement*, 29(1), 43–63. <https://doi.org/10.1080/09243453.2017.1366342>
- Friedman, R. C., & Lee, S. W. (1996). Differentiating instruction for high-achieving/gifted children in regular classrooms: A field test of three gifted-education models. *Journal for the Education of the Gifted*, 19(4), 405–436. <https://doi.org/10.1177/016235329601900403>
- Gebauer, M. M., & McElvany, N. (2017). Empirische Arbeit: Zur Bedeutsamkeit unterrichtsbezogener heterogenitätsspezifischer Einstellungen angehender Lehrkräfte für intendiertes Unterrichtsverhalten [On the significance of instructional heterogeneity-specific attitudes of prospective teachers for intended teaching behavior]. *Psychologie in Erziehung und Unterricht*, 64(3), 163–180. <https://doi.org/10.2378/peu2017.art11d>
- General Assembly of the United Nations (2019). Transforming our world: The 2030 agenda for sustainable development. In Walker, J. I., Pekmezovic, A., & Walker, G. (Eds.), *Sustainable development goals: Harnessing business to achieve the SDGs through finance, technology and law reform* (pp. 333–374). Wiley.
- *Gilson, C. M., Little, C. A., Ruegg, A. N., & Bruce-Davis, M. (2014). An investigation of elementary teachers' use of follow-up questions for students at different reading levels. *Journal of Advanced Academics*, 25(2), 101–128. <https://doi.org/10.1177/1932202X14532257>
- Gubbins, E. J. (2013). Cognitive and affective outcomes of pull-out programs. In C. M. Callahan & H. L. Hertberg-Davis (Eds.), *Fundamentals of gifted education: Considering multiple perspectives* (pp. 176–187). Routledge.

- Gutman, L. M., & Schoon, I. (2013, November 21). *The impact of non-cognitive skills on outcomes for young people: Literature review*. <https://helmtraining.co.uk/wp-content/uploads/2017/03/Gutman-and-Schoon-Impact-of-non-cognitive-skills-on-outcomes-for-young-people-1.pdf>
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341–370. <https://doi.org/10.3102/1076998606298043>
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. <https://doi.org/10.1002/jrsm.5>
- Heinle, A., Schiepe-Tiska, A., Reinhold, F., Heine, J.-H., & Holzberger, D. (2022). Supporting student motivation in class: The motivational potential of tasks. *Zeitschrift für Erziehungswissenschaft*. <https://doi.org/10.1007/s11618-022-01090-3>
- Heller, K. A., Perleth, C., & Lim, T. K. (2005). The Munich model of giftedness designed to identify and promote gifted students. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 147–170). Cambridge University Press.
- Heyvaert, M., Maes, B., & Onghena, P. (2013). *Mixed methods research synthesis: Definition, framework, and potential*. *Quality and Quantity*, 47, 659–676. <https://doi.org/10.1007/s11135-011-9538-6>
- Hughes, L. (1999). Action research and practical inquiry: How can I meet the needs of the high-ability student within my regular education classroom? *Journal for the Education of the Gifted*, 22(3), 282–297. <https://doi.org/10.1177/016235329902200304>
- *Jarvis, J., & Henderson, L. (2015). Current practices in the education of gifted and advanced learners in south Australian schools. *Australasian Journal of Gifted Education*, 24(2), 70–86. <https://doi.org/10.21505/ajge.2015.0018>
- *Johnsen, S. K., Haensly, P. A., Ryser, G. R., & Ford, R. F. (2002). Changing general education classroom practices to adapt for gifted students. *Gifted Child Quarterly*, 46(1), 45–63. <https://doi.org/10.1177/001698620204600105>
- *Kahveci, N. G., & Akgül, S. (2014). Gifted and talented students' perceptions on their schooling: A survey study. *Gifted and Talented International*, 29(1–2), 79–91. <https://doi.org/10.1080/15332276.2014.11678431>
- *Kanevsky, L. (2011). Deferential differentiation: What types of differentiation do students want? *Gifted Child Quarterly*, 55(4), 279–299. <https://doi.org/10.1177/0016986211422098>
- Kaplan, S. N. (1986). The grid: A model to construct differentiated curriculum for the gifted. In J. S. Renzulli (Ed.), *Systems and models for developing programs for the gifted and talented* (pp. 180–193). Creative Learning Press.
- *Kim, K. H., VanTassel-Baska, J., Bracken, B. A., Feng, A., & Stambaugh, T. (2014). Assessing science reasoning and conceptual understanding in the primary grades using standardized and performance-based assessments. *Journal of Advanced Academics*, 25(1), 47–66. <https://doi.org/10.1177/1932202X14520946>
- Kim, M. (2016). A meta-analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly*, 60(2), 102–116. <https://doi.org/10.1177/0016986216630607>
- *Kondor, C. A. H. (2007). *One size may not fit all, but the right teaching strategies might: The effects of differentiated instruction on the motivation of talented and gifted students* [Master Thesis, Portland State University]. ERIC. <https://files.eric.ed.gov/fulltext/ED497701.pdf>
- Kultusministerkonferenz (2009, December 10). *Grundsatzposition der Länder zur begabungsgerechten Förderung* [Basic position of the federal states on giftedness-oriented support].

- https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2009/2009_12_12-Begabungsgerechte-Foerderung.pdf
- Kultusministerkonferenz (2015, June 11). *Förderstrategie für leistungsstarke Schülerinnen und Schüler* [Support strategy for high-achieving students]. https://www.kmk.org/fileadmin/Dateien/pdf/350-KMK-TOP-011-Fu-Leistungsstarke_-_neu.pdf
- 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>
- *Laine, S., & Tirri, K. (2016). How Finnish elementary school teachers meet the needs of their gifted students. *High Ability Studies*, 27(2), 149–164. <https://doi.org/10.1080/13598139.2015.1108185>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- *Latz, A. O., Speirs Neumeister, K. L., Adams, C. M., & Pierce, R. L. (2008). Peer coaching to improve classroom differentiation: Perspectives from project CLUE. *Roeper Review*, 31(1), 27–39. <https://doi.org/10.1080/02783190802527356>
- Lee, R. P., Hart, R. I., Watson, R. M., & Rapley, T. (2015). Qualitative synthesis in practice: Some pragmatics of meta-ethnography. *Qualitative Research*, 15(3), 334–350. <https://doi.org/10.1177/1468794114524221>
- Lindner, K-T, & Schwab, S. (2020). Differentiation and individualisation in inclusive education: A systematic review and narrative synthesis. *International Journal of Inclusive Education*, 1–21. <https://doi.org/10.1080/13603116.2020.1813450>
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4), 423–458. <https://doi.org/10.3102/00346543066004423>
- *MacIntyre, H., & Ireson, J. (2002). Within-class ability grouping: Placement of pupils in groups and self-concept. *British Educational Research Journal*, 28(2), 249–263. <https://doi.org/10.1080/01411920120122176>
- Maker, C. J., & Schiever, S. W. (2010). *Curriculum development and teaching strategies for gifted learners* (3rd ed.). PRO-ED.
- Matthews, M. K. (1992). *Gifted students and whole language: A descriptive study of four classrooms* [Paper presentation]. Annual Meeting of the American Educational Research Association, San Francisco.
- *McCoach, D. B., Gubbins, E. J., Foreman, J., Rubenstein, L. D., & Rambo-Hernandez, K. E. (2014). Evaluating the efficacy of using predifferentiated and enriched mathematics curricula for grade 3 students: A multisite cluster-randomized trial. *Gifted Child Quarterly*, 58(4), 272–286. <https://doi.org/10.1177/0016986214547631>
- McCoach, D. B., & Siegle, D. (2003). Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gifted Child Quarterly*, 47(2), 144–154. <https://doi.org/10.1177/001698620304700205>
- *Meyer, B. J. F., Wijekumar, K. K., & Lin, Y-C (2011). Individualizing a web-based structure strategy intervention for fifth graders' comprehension of nonfiction. *Journal of Educational Psychology*, 103(1), 140–168. <https://doi.org/10.1037/a0021606>
- Missett, T. C., Brunner, M. M., Callahan, C. M., Moon, T. R., & Price Azano, A. (2014). Exploring teacher beliefs and use of acceleration, ability grouping, and formative assessment. *Journal for the Education of the Gifted*, 37(3), 245–268. <https://doi.org/10.1177/0162353214541326>
- Moon, S. M., Kolloff, P., Robinson, A., Dixon, F., & Feldhusen, J. F. (2009). The purdue three-stage model. In J. S. Renzulli, E. J. Gubbins, K. S. McMillen, R. D. Eckert, & C. A. Little

- (Eds.), *Systems & models for developing programs for the gifted & talented* (pp. 289–321). Prufrock Press.
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. <https://doi.org/10.1177/1094428106291059>
- National Association for Gifted Children (2010). *Redefining giftedness for a new century: Shifting the paradigm*. <https://www.nagc.org/sites/default/files/Position%20Statement/Redefining%20Giftedness%20for%20a%20New%20Century.pdf>
- Neuendorf, C., Jansen, M., Kuhl, P., & Vock, M. (2022). Wer ist leistungsstark? Operationalisierung von Leistungsstärke in der empirischen Bildungsforschung seit dem Jahr 2000 [Who is a high achiever? Operationalisation of high achievement in empirical educational research since the year 2000]. *Zeitschrift für Pädagogische Psychologie, 36*(1), 1–19. <https://doi.org/10.1024/1010-0652/a000343>
- Ninkov, I. (2020). Education policies for gifted children within a human rights paradigm: A comparative analysis. *Journal of Human Rights and Social Work, 5*(4), 280–289. <https://doi.org/10.1007/s41134-020-00133-1>
- Noetel, M., Ciarrochi, J., van Zanden, B., & Lonsdale, C. (2019). Mindfulness and acceptance approaches to sporting performance enhancement: A systematic review. *International Review of Sport and Exercise Psychology, 12*(1), 139–175. <https://doi.org/10.1080/1750984X.2017.1387803>
- *Nomi, T. (2010). The effects of within-class ability grouping on academic achievement in early elementary years. *Journal of Research on Educational Effectiveness, 3*(1), 56–92. <https://doi.org/10.1080/19345740903277601>
- OECD (2013). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs (Volume III)*. OECD Publishing. <https://doi.org/10.1787/9789264201170-en>
- OECD (2019). *PISA 2018 Assessment and analytical framework*. OECD Publishing.
- *Oswald, M., & Villiers, J. M. D. (2013). Including the gifted learner: Perceptions of South African teachers and principals. *South African Journal of Education, 33*(1), 1–21. <https://doi.org/10.15700/saje.v33n1a603>
- *Parish, L. (2016). The power of creativity: A case-study of a mathematically highly capable grade 5 student. In B. White, M. Chinnappan, & S. Trenholm (Eds.), *Proceedings of the 39th annual conference of the mathematics education research group of Australasia, opening up mathematics education research* (pp. 511–518). MERGA.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing.
- Philippe, D. L. (2017). *Implementing social emotional learning (SEL): An evaluation of Illinois teachers' capacity to provide SEL instruction and use the Illinois SEL Standards* [Doctoral dissertation, Loyola University Chicago]. Loyola eCommons. https://ecommons.luc.edu/luc_diss/2839/
- *Prast, E. J., & de van Weijer-Bergsma, 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. (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>
- Puzio, K., Colby, G. T., & Algeo-Nichols, D. (2020). Differentiated literacy instruction: Boondoggle or best practice? *Review of Educational Research, 90*(4), 459–498. <https://doi.org/10.3102/0034654320933536>

- Puzio, K., Newcomer, S. N., & Goff, P. (2015). Supporting literacy differentiation: Principal's role in a community of practice. *Literacy Research and Instruction, 54*(2), 135–162. <https://doi.org/10.1080/19388071.2014.997944>
- *Reis, S. M., Gubbins, E. J., Briggs, C. J., Schreiber, F. J., Richards, S., Jacobs, J. K., Eckert, R. D., & Renzulli, J. S. (2004). Reading instruction for talented readers: Case studies documenting few opportunities for continuous progress. *Gifted Child Quarterly, 48*(4), 315–338. <https://doi.org/10.1177/001698620404800406>
- *Reis, S. M., Little, C. A., Fogarty, E., Housand, A. M., Housand, B. C., Sweeny, S. M., & Muller, L. M. (2010). *Case studies of successful schoolwide enrichment model-reading (SEM-R) classroom implementations*. National Research Center on the Gifted and Talented.
- Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Differences, 20*(4), 308–317. <https://doi.org/10.1016/j.lindif.2009.10.012>
- Rinn, A. N., Mun, R. U., & Hodges, J. (2020). *2018-2019 State of the states in gifted education*. National Association for Gifted Children and the Council of State Directors of Programs for the Gifted. <https://www.nagc.org/2018-2019-state-states-gifted-education>
- *Robinson, A., Dailey, D., Hughes, G., & Cotabish, A. (2014). The effects of a science-focused STEM intervention on gifted elementary students' science knowledge and skills. *Journal of Advanced Academics, 25*(3), 189–213. <https://doi.org/10.1177/1932202X14533799>
- *Rubenstein, L. D., Gilson, C. M., Bruce-Davis, M. N., & Gubbins, E. J. (2015). Teachers' reactions to pre-differentiated and enriched mathematics curricula. *Journal for the Education of the Gifted, 38*(2), 141–168. <https://doi.org/10.1177/0162353215578280>
- *Russell, J. L. (2018). High school teachers' perceptions of giftedness, gifted education, and talent development. *Journal of Advanced Academics, 29*(4), 275–303. <https://doi.org/10.1177/1932202X18775658>
- *Saleh, M., Lazonder, A. W., & Jong, T. de. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science, 33*(2), 105–119. <https://doi.org/10.1007/s11251-004-6405-z>
- Schleicher, A. (2016). *Teaching excellence through professional learning and policy reform: Lessons from around the world. International summit on the teaching profession*. OECD Publishing.
- *Seedorf, S. (2014). Response to intervention: Teachers' needs for implementation in gifted and talented programs. *Gifted Child Today, 37*(4), 248–257. <https://doi.org/10.1177/1076217514544029>
- *Shaunessy-Dedrick, E., Evans, L., Ferron, J., & Lindo, M. (2015). Effects of differentiated reading on elementary students' reading comprehension and attitudes toward reading. *Gifted Child Quarterly, 59*(2), 91–107. <https://doi.org/10.1177/0016986214568718>
- Shayshon, B., Gal, H., Tesler, B., & Ko, E-S (2014). Teaching mathematically talented students: A cross-cultural study about their teachers' views. *Educational Studies in Mathematics, 87*(3), 409–438. <https://doi.org/10.1007/s10649-014-9568-9>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- *Simmons, A. (2018). *Second grade teachers' perspectives on differentiated instruction* [Doctoral dissertation, Walden University]. Walden Dissertations and Doctoral Studies. <https://scholarworks.waldenu.edu/dissertations/5591>.

- Singer, L. M., & Alexander, P. A. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*(6), 1007–1041. <https://doi.org/10.3102/0034654317722961>
- Sisk, D. (2009). Myth 13: The regular classroom teacher can 'go it alone'. *Gifted Child Quarterly, 53*(4), 269–271. <https://doi.org/10.1177/0016986209346939>
- 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*, 1–23. <https://doi.org/10.3389/fpsyg.2019.02366>
- Sparfeldt, J. R., Schilling, S. R., & Rost, D. H. (2004). Segregation oder Integration: Einstellungen potenziell Betroffener zu Fördermaßnahmen für hochbegabte Jugendliche [Integration or separation? Attitudes toward nurturing strategies for gifted adolescents]. *Reportpsychologie, 29*(3), 170–176.
- Steenbergen-Hu, S., & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly, 55*(1), 39–53. <https://doi.org/10.1177/0016986210383155>
- Stoeger, H. (2009). The history of giftedness research. In L. V. Shavinina (Ed.), *International handbook on giftedness: Part one* (pp. 17–38). Springer Netherlands.
- Stohr-Hunt, P. M. (1996). An analysis of frequency of hands-on experience and science achievement. *Journal of Research in Science Teaching, 33*(1), 101–109. [https://doi.org/10.1002/\(SICI\)1098-2736\(199601\)33:1 <101::AID-TEA6>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1098-2736(199601)33:1 <101::AID-TEA6>3.0.CO;2-Z)
- Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal, 7*(7), 935–947.
- *Swiatek, M. A., & Lupkowski-Shoplik, A. (2003). Elementary and middle school student participation in gifted programs: Are gifted students underserved? *Gifted Child Quarterly, 47*(2), 118–130. <https://doi.org/10.1177/001698620304700203>
- *Tieso, C. (2005). The effects of grouping practices and curricular adjustments on achievement. *Journal for the Education of the Gifted, 29*(1), 60–89. <https://doi.org/10.1177/016235320502900104>
- Tod, D., Booth, A., & Smith, B. (2022). Critical appraisal. *International Review of Sport and Exercise Psychology, 15*(1), 52–72. <https://doi.org/10.1080/1750984X.2021.1952471>
- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Association for Supervision & Curriculum Development.
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.
- Tomlinson, C. A., Brighton, C. M., 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/10.1177/016235320302700203>
- Trivette, C. M., Dunst, C. J., Hamby, D. W., & Meter, D. (2012). Relationship between early childhood practitioner beliefs and the adoption of innovative and recommended practices. *Research Brief, 6*(1), 1–12.
- Troxclair, D. A. (2000). Differentiating instruction for gifted students in regular education social studies classes. *Roeper Review, 22*(3), 195–198. <https://doi.org/10.1080/02783190009554033>
- van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review, 28*(100281), 1–18. <https://doi.org/10.1016/j.edurev.2019.05.003>

- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Viechtbauer, W. (2021). *Morris (2008)*. The metafor package: A meta-analysis package for R. <https://www.metafor-project.org/doku.php/analyses:morris2008>
- *Vreys, C., Ndungbogun, G. N., Kieboom, T., & Venderickx, K. (2018). Training effects on Belgian preschool and primary school teachers' attitudes towards the best practices for gifted children. *High Ability Studies*, 29(1), 3–22. <https://doi.org/10.1080/13598139.2017.1312295>
- Vygotsky, L. S. (1978). Interaction between learning and development. In M. Gauvain & M. Cole (Eds.), *Readings on the development of children* (pp. 33–40). Scientific American Books.
- Wendt, H., Bos, W., Goy, M., & Jusufi, D. (2017). *TIMSS 2015: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen* [TIMSS 2015: Scale manual for documenting the survey instruments and working with the data sets]. Waxmann.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(2), 12. <https://doi.org/10.14507/epaa.v10n12.2002>
- Westberg, K. L., Archambault, F. X., Dobyns, S. M., & Salvin, T. J. (1993). The classroom practices observation study. *Journal for the Education of the Gifted*, 16(2), 120–146. <https://doi.org/10.1177/016235329301600204>
- White, K., Fletcher-Campbell, F., & Ridley, K. (2003). *What works for gifted and talented pupils: A review of recent research*. National Foundation for Educational Research.
- Williams, F. E. (1986). The cognitive-affective interaction model for enriching gifted programs. In J. S. Renzulli (Ed.), *Systems and models for developing programs for the gifted and talented* (pp. 461–484). Creative Learning Press.
- Wirtz, M., & Kutschmann, M. (2007). Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Masse [Analyzing interrater agreement for categorical data using Cohen's kappa and alternative coefficients]. *Die Rehabilitation*, 46(6), 370–377. <https://doi.org/10.1055/s-2007-976535>
- *Ysseldyke, J., & Tardrew, S. (2007). Use of a progress monitoring system to enable teachers to differentiate mathematics instruction. *Journal of Applied School Psychology*, 24(1), 1–28. https://doi.org/10.1300/J370v24n01_01
- *Ysseldyke, J., Tardrew, S., Betts, J., Thill, T., & Hannigan, E. (2004). Use of an instructional management system to enhance math instruction of gifted and talented students. *Journal for the Education of the Gifted*, 27(4), 293–310. <https://doi.org/10.4219/jeg-2004-319>
- *Yuen, M., Chan, S., Chan, C., Fung, D. C. L., Cheung, W. M., Kwan, T., & Leung, F. K. S. (2018). Differentiation in key learning areas for gifted students in regular classes. *Gifted Education International*, 34(1), 36–46. <https://doi.org/10.1177/0261429416649047>

About the Authors

Lisa Ziernwald is a research associate at the professorship for Learning and Instruction at the TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment (ZIB), Germany. Her research interests focus on high-achieving students, digital learning, research syntheses, and PISA.

Delia Hillmayr is a research associate at the professorship for Learning and Instruction at the TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment (ZIB), Germany. Her research interests focus on learning with digital tools, dealing with heterogeneity in educational contexts, and preparing evidence for educational practice.

Doris Holzberger is a professor for Learning and Instruction at TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment, Germany. Her research interests focus on school characteristics as predictors for students' development, instructional quality, and characteristics of successful teachers.