Technische Universität München
TUM School of Social Sciences and Technology

Classification and Promotion of High-Achieving Students:

A Multidimensional Perspective on their Characteristics and Needs in Inclusive Education

Lisa Marie Ziernwald

Vollständiger Abdruck der von der TUM School of Social Sciences and Technology der Technischen Universität München zur Erlangung des akademischen Grades einer **Doktorin der Philosophie (Dr. phil)** genehmigten Dissertation.

Vorsitz: Prof. Dr. Doris Holzberger

Prüferinnen der Dissertation:

1. Prof. Dr. Kristina Reiss
2. Prof. Dr. Heidrun Stöger
3. Prof. Dr. Doris Lewalter-Manhart

Die Dissertation wurde am 16.11.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Social Sciences and Technology am 07.02.2023 angenommen.

Acknowledgments

Writing my dissertation not only led to obtaining an academic degree but also contributed a great deal to my personal development. I am deeply grateful to all the people who have been with me through both the academic and the personal ups and downs during my doctoral journey.

Firstly, I would like to thank my doctoral supervisor Professor Dr. Kristina Reiss, who helped me to initiate, conduct, and finish this journey with all the means at her disposal. Right at the beginning, she told me that her doctorate was challenging but also one of the best times in her life. One has the possibility to work on a topic you love in great detail while experiencing various other interesting and instructive aspects, which broaden your horizon. Now, at the end of my doctorate, I fully agree with this statement and express my deepest gratitude to Professor Dr. Kristina Reiss for making this possible. My thanks go further to Professor Dr. Heidrun Stöger, who supported me with her enormous expertise in giftedness research and Professor Dr. Doris Lewalter-Manhart, who supported me and my work in the PISA team.

Further, Professor Dr. Anja Schiepe-Tiska was a crucial companion for me during my doctoral journey. In addition to various professional discussions about research, I highly appreciated the personal conversations and the insights into areas of research that were previously unknown to me. Anja told me that there rarely is a right or wrong way in research, but that the situation is more complex and that you have to think about various ways and then decide on the most appropriate one.

On my doctoral journey, I repeatedly learned that sometimes the little things in life are those which contribute to big changes. Therefore, I want to thank all my colleagues, especially those from the Chair for Research on Learning and Instruction and the PISA team, for the enormous support. I enjoyed the critical discussions about research as well as the joint lunch breaks, team events, and conference attendances. In this context, I also want to especially thank Professor Dr. Doris Holzberger, who always trusted me, gave me the freedom to work independently on different projects, and at the same time was always available when I needed advice or support. I also thank her very much for chairing the examination committee.

Finally, from the bottom of my heart, I would like to thank all my family and friends who helped me find a good balance between my doctoral journey and the other wonderful parts of my life. Even though I was sometimes close to losing faith in successfully completing my doctoral journey, you always supported me and gave me the strength to go a long way, and brought me to where I am today. I would never again miss any of those experiences.

Many thanks to each and every one of you!

Abstract

High-achieving students, who can be seen as one essential resource of society, have moved into focus and became part of the discussion about educational promotion approaches. Thereby two problems occur. First, high-achieving students are frequently considered as one homogeneous group and second, their promotion is often neglected. Hence, this research project aims to investigate the extent and quality of high-achieving students' heterogeneity and how it can be identified from a multidimensional perspective, including different achievement indicators and motivational-affective characteristics. Further, it is investigated how high-achieving students' heterogeneity can be considered to offer the best possible support in learning through differentiated instruction (DI) in an inclusive educational setting.

To investigate high-achieving students' heterogeneity from a multidimensional perspective and thus improve their classification, domain-specific (i.e., mathematical competences) as well as domain-general (i.e., figural reasoning) achievement indicators were used to identify high-achieving student subgroups (Paper A). In addition, domain-specific and domain-general motivational-affective characteristics were used to further characterize the detected high-achieving student subgroups. From a methodological perspective, traditional giftedness research[1] often used the cut-off score approach, whereas the latent profile analysis approach has recently gained popularity. To investigate if these two approaches can be used interchangeably, both approaches with the same sample of high-achieving students (*N* = 1563) were used and the respective classification results were compared. Results showed a remarkable achievement and motivational-affective heterogeneity between high-achieving student subgroups. Both methodological approaches resulted in the four largest detected subgroups of high-achieving students on the aggregated level (i.e., Top Performer, Overachiever, Underachiever, Operator), but could not be used interchangeably on an individual level (assignment of the individual high-achieving student to subgroups). Hence, the need for a multidimensional perspective on high-achieving students considering different achievement indicators and motivational-affective characteristics is emphasized.

Further, the promotion of high-achieving students in inclusive educational settings through DI was investigated (Paper B). Therefore, a systematic mixed-methods review (*N* = 49 studies) was used to summarize the impact of DI on high-achieving students' outcomes as well as different conditions for the successful implementation of DI in mixed-ability classrooms. Results indicated that DI has a mainly positive impact on high-achieving students' outcomes and is therefore a promising approach for their promotion in inclusive educational settings. However, DI was not used proactively and regularly for high-achieving students, despite teachers and high-achieving students having affirmed its usefulness. Barriers (e.g., teachers' misconceptions) and facilitators (e.g., professional staff development) that might hinder or help to implement DI successfully are discussed.

To sum up, this dissertation contributes, from a multidimensional perspective, to a better understanding of high-achieving students' heterogeneity. Further, the importance of individualized promotion in inclusive educational settings to fully develop their potential is underlined.

[1] In this dissertation, the German term *Begabtenforschung*, focusing on the research of high-achieving students, is translated with *giftedness research* and therefore includes next to gifted students also high-achieving students. The same is true for gifted education.

Zusammenfassung

Leistungsstarke Schüler*innen, die als eine wesentliche Ressource der Gesellschaft angesehen werden können, rückten in den letzten Jahren in den Fokus und wurden Teil der Diskussion über Förderansätze im Bildungssystem. Dabei treten zwei Probleme auf. Erstens werden leistungsstarke Schüler*innen häufig als eine homogene Gruppe betrachtet und zweitens wird ihre Förderung häufig vernachlässigt. Ziel der vorliegenden Forschungsarbeit ist es daher, das Ausmaß und die Qualität der Heterogenität von leistungsstarken Schüler*innen aus einer mehrdimensionalen Perspektive zu betrachten. Diese berücksichtigt neben verschiedenen Leistungsindikatoren auch motivational-affektive Merkmale. Darüber hinaus wird untersucht, wie die Heterogenität leistungsstarker Schüler*innen in einem inklusiven Bildungskontext durch differenzierten Unterricht (DI) berücksichtigt werden kann, um eine bestmögliche Unterstützung beim Lernen zu bieten.

Um die Heterogenität leistungsstarker Schüler*innen aus einer mehrdimensionalen Perspektive zu untersuchen und damit deren Klassifikation zu verbessern, wurden sowohl domänenspezifische (mathematische Kompetenzen) als auch domänenübergreifende Leistungsindikatoren (figurales schlussfolgerndes Denken) verwendet, um Subgruppen leistungsstarker Schüler*innen zu identifizieren (Paper A). Weiter wurden domänenspezifische und domänenübergreifende motivational-affektive Merkmale verwendet, um die ermittelten Subgruppen leistungsstarker Schüler*innen genauer zu charakterisieren. Aus methodischer Sicht wurde in der traditionellen Begabtenforschung häufig der Cut-Off Score Ansatz verwendet, während in letzter Zeit der Ansatz der latenten Profilanalyse an Bedeutung gewonnen hat. Um zu untersuchen, ob diese beiden Ansätze auswechselbar eingesetzt werden können, wurden beide Ansätze mit der gleichen Stichprobe leistungsstarker Schüler*innen (N = 1563) angewandt und die jeweiligen Klassifikationsergebnisse verglichen. Es zeigte sich eine bemerkenswerte Heterogenität zwischen den gefundenen Subgruppen der leistungsstarken Schüler*innen bezüglich ihrer Leistung und den motivational-affektiven Eigenschaften. Beide methodischen Ansätze ergaben auf aggregierter Ebene die gleichen vier größten Subgruppen von leistungsstarken Schüler*innen (Top Performer, Overachiever, Underachiever, Operator). Auf individueller Ebene (Zuordnung einzelner leistungsstarker Schüler*innen zu Subgruppen) konnten die Ansätze nicht austauschbar verwendet werden. Dementsprechend wird die Notwendigkeit einer mehrdimensionalen Perspektive auf leistungsstarke Schüler*innen unter Berücksichtigung verschiedener Leistungsindikatoren und motivational-affektiver Merkmale betont.

Des Weiteren wurde untersucht, wie DI zur Förderung leistungsstarker Schüler*innen in einem inklusiven Bildungskontext eingesetzt werden kann (Paper B). Dafür wurden in einem systematischen mixed-methods Review (N = 49 Studien) die Auswirkungen von DI auf die Ergebnisse leistungsstarker Schüler*innen sowie die verschiedenen Bedingungen für eine erfolgreiche Umsetzung von DI für leistungsstarke Schüler*innen in leistungsheterogenen Klassen zusammenfassend analysiert. Die Ergebnisse deuten darauf hin, dass DI eine überwiegend positive Auswirkung auf die Ergebnisse leistungsstarker Schüler*innen hat und daher ein vielversprechender Ansatz für deren Förderung in inklusiven Bildungskontexten ist. Allerdings wurde DI nicht proaktiv und regelmäßig für leistungsstarke Schüler*innen eingesetzt, obwohl Lehrkräfte und leistungsstarke Schüler*innen dessen Nützlichkeit bestätigten. Es werden Hin-

dernisse (z. B. Fehlvorstellungen der Lehrkräfte) und Unterstützungsmöglichkeiten (z. B. Lehrkräftefortbildungen) diskutiert, die eine erfolgreiche Umsetzung von DI behindern beziehungsweise unterstützen können.

Zusammenfassend trägt die vorliegende Dissertation aus einer mehrdimensionalen Perspektive zu einem besseren Verständnis der Heterogenität leistungsstarker Schüler*innen bei. Außerdem wird die Bedeutung einer individualisierten Förderung in inklusiven Bildungskontexten unterstrichen, um deren Potenzial vollständig zu entfalten.

Table of Contents

Acknowledgments i
Abstract ii
Zusammenfassung iii
1 Introduction 1
2 Identification of High-Achieving Student Subgroups 4
2.1 Multidimensional Classification and Characterization of High-Achieving Students 4
2.2 Methodological Classification Approaches 6
2.2.1 Cut-Off Score Approach 6
2.2.2 Latent Profile Analyses 8
2.3 Summary 9
3 Promotion of High-Achieving Students 10
3.1 The Role of High-Achieving Students in the Educational Context 10
3.2 Enrichment, Acceleration, and Grouping 11
3.3 Differentiated Instruction in Mixed-Ability Classrooms 12
3.4 Summary 14
4 The Present Research 16
5 Summary of Publications 19
5.1 Identification and Characterization of High-Achieving Student Subgroups (Paper A) 19
5.1.1 Research Aims 19
5.1.2 Methodology 20
5.1.3 Main Results 21
5.2 Promotion of High-Achieving Students through Differentiated Instruction (Paper B) 21
5.2.1 Research Aims 22
5.2.2 Methodology 22
5.2.3 Main Results 23
6 Discussion 24
6.1 Discussion of the Main Findings 24
6.2 Classification of High-Achieving Students – Is it Really Necessary? 26
6.3 Gifted Education—Praise or Malice for the Education System 27
6.4 Methodological Considerations 29
6.4.1 Methodological Strengths 29
6.4.2 Limitations and Directions for Future Research 30
6.5 Practical Implications 32
7 Conclusion 35
References 36

Appendix 50
Appendix A - Paper A 50
Appendix B - Paper B 68

1 Introduction

“No two children are equal. As Aristotle—and, later, Thomas Jefferson—observed so astutely, ‘Nothing is so unequal as the equal treatment of unequals’.”
(Cooper, 2009, p. 285)

The appropriate education of high-achieving students is an important educational goal in order to meet their needs and develop their full potential. The promotion of high-achieving students serves the individual student as well as the society as a whole (Plucker & Callahan, 2020) and has gained importance in recent years (Kultusministerkonferenz [KMK], 2015, 2016a; Moltzen, 2011). Thereby, two main problems related to the introductory quote appear. First, all high-achieving students have often been considered as equal and therefore as one homogeneous group. Second, promotion for high-achieving students was often neglected in the educational context and their individual differences were not appropriately taken into account in order to develop their full potential.

With regard to the appropriate education of high-achieving students, one of the most dangerous and incorrect myths in gifted education is to consider high-achieving students as one homogeneous group (Reis & Renzulli, 2009). In contrast, high-achieving students are a multifaceted group with strengths and weaknesses in different achievement and motivational-affective areas (Tirri & Laine, 2017). Consequently, the belief that high-achieving students are competent in all different areas was repeatedly questioned (Sak, 2011; Tirri & Laine, 2017) and a multidimensional perspective on high-achieving students including motivational-affective characteristics besides different achievement indicators was established (e.g. Heller et al., 2005; Stoeger, 2009). In order to consider high-achieving students’ heterogeneity, a continuum with two ends is conceivable. At one end, high-achieving students can be considered as one homogeneous group and at the other end, each high-achieving student can be treated individually. From research, we know that considering high-achieving students as one homogeneous group is not reasonable (e.g., Castejón et al., 2016; Hofer & Stern, 2016; Leikin et al., 2017). Although from a statistical perspective, we can treat each high-achieving student individually (i.e., person-specific approach; Howard & Hoffman, 2018), it is very tedious and hardly implementable from a practical perspective in daily lessons. Therefore, despite the idealism that all students are educated according to their individual needs, we must not forget the aspect of feasibility in the everyday school context (Trautmann & Wischer, 2011).

Hence, the classification of high-achieving students into subgroups through person-centered approaches appear as a suitable compromise between treating all high-achieving students as one group and the ideal world, where all high-achieving students would be considered individually with their respective strengths and weaknesses (Howard & Hoffman, 2018). Person-centered classification approaches aim to detect subgroups of individuals that show comparable patterns of certain characteristics. Thereby, the goal of classification is to define subgroups as parsimony as possible but as complex as necessary to appropriately describe the patterns of the underlying population (Collins & Lanza, 2010). So far, research using a multidimensional perspective on high-achieving students is still sparse (Hong & Aqui, 2004) and the empirical investigation of the heterogeneity across different high-achieving student subgroups is necessary (Castejón et al., 2016).

At the individual level, an important educational goal is to provide all students, including the high-achieving ones, support measures according to their individual abilities so that all students can develop their full potential (Dixson et al., 2020; KMK, 2009; Lucito, 1964). If high-achieving students are challenged too little for longer periods, it can have numerous negative consequences (e.g., decline in performance and motivation; KMK, 2015). In order to improve teaching and adapt instruction to the different strengths and weaknesses of high-achieving students and to provide individualized promotion, teachers need to know how high-achieving students can be classified and which characteristics they have (Castejón et al., 2016; KMK, 2015; Lazarides et al., 2019).

At the society level, the appropriate development of high-achieving students is relevant, as they can be seen as "the most important human resource for a nation and the world" (Kahveci & Akgül, 2014, p. 87). People taking responsibility and guidance regarding global challenges in the world as well as highly skilled professionals on the labor market are needed (Cappelli, 2015; European Centre for the Development of Vocational Training, 2016; OECD, 2020). Thus, there is high demand for high-achieving students in today's societies (Ziegler & Stoeger, 2012). Thereby, the interest lies not only in highly gifted students who may become future mathematics or science experts, but also in broadening the scope to those students who understand the social implications that mathematics plays in the world and who are able to fulfill an enlightened role in their actions. This approach aligns with current global and political visions and movements to enable a broader student body to take on responsibility and leadership for the world's global challenges (KMK, 2016a; OECD, 2020). A society's most important human resource might be ruined if high-achieving students are not challenged appropriately. In contrast, education that considers high-achieving students' needs has the potential to increase their quality of life (Tirri & Laine, 2017).

In addition, aiming for individualized promotion gained particular importance as inclusive education became more popular worldwide (Armstrong, 2011; Lindner & Schwab, 2020). The guiding principle of inclusive education is to adapt the general education system such that all students can participate in general educational institutions (Gebhardt & Heimlich, 2018). Inclusive education can therefore be defined as "nondiscriminatory education for all" (Tirri & Laine, 2017, p. 239; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2009) aiming to equally meet the needs of high-achieving, struggling, and all other students (Armstrong & Richards, 2011; Booth & Ainscow, 2002). The assumption that inclusive education also embraces high-achieving students has not always been supported and even today the focus of inclusive education is often on struggling students, as for example, students with disabilities (e.g., Amor et al., 2019; Krämer et al., 2021; Tirri & Laine, 2017). Nevertheless, the importance of also considering high-achieving students' needs in inclusive education is further underlined as most high-achieving students are educated in mixed-ability classrooms (e.g., White et al., 2003; Yuen et al., 2018) and not in self-contained gifted programs.

Without adapting instruction to the individual needs of high-achieving students, they are "the most ill-served" student group (Borland, 2005, p. 2). Hence, with the trend towards inclusive education it becomes even more apparent that the one-size-fits-all approach isn't appropriate and the question how the different prerequisites and needs of all the diverse learners, including high-achieving students, can be taken into account, arises (Gebhardt & Heimlich, 2018;

Lindner & Schwab, 2020). One promising teaching approach in an inclusive educational setting is differentiated instruction (DI), which builds on students commonalities and at the same time considers students' differences to provide all students the possibility to develop their potential effectively (Tomlinson, 2001). Accordingly, one quality criterion of inclusive education is differentiation that does not focus on single students but considers all students within an inclusive educational setting (Arndt et al., 2014). Positive effects of DI on students in general were already found (Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019). However, it remains an open question if this positive impact can also be generalized for high-achieving students and which facilitators and barriers might influence its successful implementation for high-achieving students in inclusive educational settings.

To sum up, a crucial open question in giftedness research and gifted education is how the multidimensional characteristics of high-achieving students can be classified and appropriately taken into account through differentiated promotion in inclusive educational settings. This dissertation contributes for the long-neglected group of high-achieving students important theoretical, methodological, and practical implications regarding this open question through the consideration of four important gifted education topics (cf. Davis, 2009): (1) classification of high-achieving students, (2) characterization of high-achieving student subgroups, (3) promotion of high-achieving students in schools, and (4) individualization strategies for high-achieving students. These four research areas and the respective contributions of this dissertation are discussed in detail.

2 Identification of High-Achieving Student Subgroups

The definition and thus also the identification of high-achieving students have been widely discussed and considered an important topic in giftedness research for about 100 years (McBee et al., 2014). The identification of high-achieving students is the foundation for other highly relevant decisions, such as the inclusion and exclusion of students in promotion programs (Acar et al., 2016). However, how these students are defined, which identification indicators are used, and which methodological approach is applied varies widely between studies. This lack of comparability leads to less generalizable and hardly contrastable results across studies (Acar et al., 2016; Carman, 2013; Neuendorf et al., 2022). Each resulting sample through the use of these various achievement indicators and methodological approaches can be seen as one certain high-achieving student subgroup from the whole high-achieving student population. Thereby, a conceptual distinction between gifted and high-achieving students is frequently made. Gifted students are often mainly defined through their high domain-general potential (e.g., fluid intelligence), whereas high-achieving students are often mainly defined through their high domain-specific achievement (e.g., mathematical competence). However, despite this conceptual distinction, a large overlap between the operationalization of gifted and high-achieving students was found in research studies (e.g., Carman, 2013; Neuendorf et al., 2022). Further, multidimensional giftedness models often consider domain-general as well as domain-specific achievement (Subotnik et al., 2011, 2019). A broad understanding of high-achieving students, including gifted as well as potentially high-achieving students, is helpful for solving societal problems (KMK, 2016a; OECD, 2020). Accordingly, recent developments in giftedness research call for a more inclusive perspective by reducing the emphasis on the label *gifted* and focus more strongly on the different characteristics and needs of this student group in order to maximize learning (Borland, 2005; Dixson et al., 2020). To investigate the different characteristics of high-achieving students from a multidimensional perspective, the consideration of high-achieving student subgroups might be a good balance between the two extreme poles that consider all individuals on their own versus all high-achieving students as equal.

To sum up, in this dissertation, gifted students are seen as one part of high-achieving students and research investigating gifted as well as high-achieving students is considered. Hence, high achievement includes domain-general as well as domain-specific achievement indicators. In the following subchapters, the heterogeneity in used identification indicators and methodological classification approaches is discussed from a multidimensional perspective on high-achieving students.

2.1 Multidimensional Classification and Characterization of High-Achieving Students

Looking at giftedness research in recent decades, it can be seen that there has been a shift away from a unidimensional perspective to a multidimensional perspective on high-achieving students. In addition to different achievement indicators, this multidimensional perspective also includes motivational-affective characteristics (García-Martínez et al., 2021; Heller et al., 2005; Renzulli, 2005; Stoeger, 2009). One commonality across the different multidimensional models and the respective research studies is that they still include high intelligence as one aspect of high-achieving students. However, the operationalization of high intelligence differs widely and many studies do not mention how intelligence was measured. Next to intelligence, domain-specific achievement is often used to define high-achieving students (Carman, 2013;

Stoeger, 2009). Domain-specific indicators are relevant for high-achieving students, as their teaching is one main task of schooling (Alexander et al., 1991), and persons with high certain domain-specific achievement are needed in the labor market (Cappelli, 2015; European Centre for the Development of Vocational Training, 2016). Accordingly, some top-ranking researchers claim that support measures for high-achieving students should be domain-specific (e.g., Dixson et al., 2020).

One of the most influential theories that distinguished domain-general from domain-specific achievement in the context of intelligence is the Theory of Crystallized and Fluid Intelligence (Cattell, 1943; Stoeger, 2009). Thereby, domain-general achievement can be equated with fluid intelligence defined as “purely general ability to discriminate and perceive relations” whereas domain-specific achievement can be equated with crystallized intelligence defined as “discriminatory habits long established in a particular field” (Cattell, 1943, p. 178). In this dissertation, intelligence is defined as fluid intelligence that can be distinguished from domain-specific achievement (Baumert et al., 2007). A common distinction within domain-specific achievement is the differentiation between literacy-oriented achievement, focusing on the application of knowledge in everyday contexts and curricular-oriented achievement, focusing on knowledge taught at school. There is an ongoing discussion on whether these two achievement areas are independent (Neidorf et al., 2006; Saß et al., 2017). Despite the existence of different multidimensional models and theories (e.g., Cattell, 1943; Heller et al., 2005; Saß et al., 2017), empirical research referring to it across different high-achieving student subgroups is expandable (Hong & Aqui, 2004).

The multidimensional perspective and the empirical investigation is especially relevant in the field of high-achieving students, as a frequent assumption is that high-achieving students show their potential in all areas and therefore one indicator is enough to identify high-achieving students (Sak, 2011; Tirri & Laine, 2017; Worrell, 2009). In contrast, research has shown that high-achieving students partially show remarkable heterogeneity across different achievement indicators and therefore high-achieving student subgroups with strengths in different areas can be detected (e.g., Castejón et al., 2016; Hofer & Stern, 2016; Leikin et al., 2017; Lohman et al., 2008). Therefore, it is necessary to empirically investigate how domain-general and domain-specific achievement is related to each other within high-achieving student subgroups and how those achievement indicators are related to motivational-affective characteristics.

Next to the consideration of different achievement indicators and the cognitive perspective on high-achieving students, motivational-affective characteristics are discussed as part of the multidimensional models. Often, they are considered as moderating characteristics between domain-general potential and domain-specific manifest achievement (e.g., Heller et al., 2005). In accordance with the inclusion of motivational-affective characteristics in multidimensional giftedness models, these characteristics also gained more attention in the whole educational context as they are related to students’ success in school and at work (Gutman & Schoon, 2013; Schiepe-Tiska, 2019). Thereby, on the one hand, high-achieving students on average showed more positive motivational-affective characteristics compared to their not-as-high-achieving identified peers (e.g., Agaliotis & Kalyva, 2019; Gubbels et al., 2018). On the other hand, differences were also found between different subgroups of high-achieving students regarding their motivational-affective characteristics (e.g., Agaliotis & Kalyva, 2019; Castejón et al., 2016). The particular importance of motivational-affective characteristics for high-achieving students is further underlined through the result that the relationship between those characteristics with achievement is, on average, even stronger for high-achieving students compared to average- and low-achieving students (OECD, 2013b). Accordingly, motivational-affective characteristics are used to explain why or why not high-achieving students' potential transfers into manifest achievement and if high-achieving students' potential ideally develops (e.g., Cross & Cross, 2017; Gagné, 2004; Heller et al., 2005). However, despite the conceptual development towards a multidimensional perspective, research comparing high-achieving student subgroups regarding their achievement and motivational-affective characteristics is still sparse (Hong & Aqui, 2004).

This research project contributes to this important giftedness research field by identifying high-achieving student subgroups through domain-general and domain-specific achievement indicators and further characterizing these detected subgroups through domain-general and domain-specific motivational-affective characteristics. As a result, on the one hand for educational practice, knowledge regarding the heterogeneity of the achievement and motivational-affective characteristics across high-achieving student subgroups can be gained in order to adapt teaching. On the other hand, knowledge regarding the relationship between motivational-affective characteristics with achievement indicators is gained for theoretical multidimensional models and respective research.

2.2 Methodological Classification Approaches

The consideration of different indicators for the classification of high-achieving students came along with different methodological approaches that identified high-achieving student subgroups in order to describe the heterogeneity and patterns across the different indicators used (e.g., Castejón et al., 2016; Leikin et al., 2017; MacCallum et al., 2002; Mammadov et al., 2016). Accordingly, a shift towards the investigation of high-achieving student subgroups instead of the whole population of high-achieving students as one homogeneous group can be observed (Stoeger, 2009). Two often used methodological classification approaches are the cut-off score and the latent profile analysis (LPA) approach.

2.2.1 Cut-Off Score Approach

One commonly used approach is the dichotomization of a continuous variable resulting in two subgroups with students scoring either above or below the chosen cut-off score. Overall, different values for the cut-off score (e.g., median, one standard deviation above the mean, a beforehand fixed point) are considered and used (MacCallum et al., 2002). So far, there is no consensus which cut-off score separates high-achieving from not as high-achieving identified students (Neuendorf et al., 2022). One distinction that can be made with cut-off scores is the differentiation between cut-off scores based on norm-referenced versus criterion-referenced assessments. In criterion-referenced assessments, a student's achievement is compared to a predetermined, absolute standard independent of the achievement of other students. With a criterion-referenced assessment, one can describe which specific competences on a continuum a student has attained. In norm-referenced assessments, "a predetermined percentage of students […] would obtain a certain grade", and therefore the achievement is compared to a relative standard. With a norm-referenced assessment, one can describe if a student achieves higher or lower than other students but does not necessarily tell us what a student can and cannot do. Neither approach is ideal in extreme situations (Bond, 1996; Glaser, 1963; Lok et al., 2016, p. 450; Sternberg et al., 2022).

As giftedness is still often equated with high intelligence (Stoeger, 2009), one well-known, norm-referenced cut-off score is an IQ of 130 (e.g., Leikin et al., 2017). More than half of the studies, which reported using a cut-off score for identifying gifted students, used an IQ of 130. In contrast, another 22.5% used an IQ of 120 and, therefore, a more liberal criterion (Carman, 2013). Another norm-referenced cut-off score recommended by the National Association for Gifted Children is an achievement score in the top 10% in one specific or different achievement domains (National Association for Gifted Children, 2010). Ten percent was also the median, when cut-off scores for the identification of high-achieving students were summarized (Neuendorf et al., 2022). On the other hand, the Programme for International Student Assessment (PISA) uses a criterion-referenced cut-off score and defines high-achieving students as students scoring at one of the two highest competence levels (i.e., level five or six). How many students reach this competence level differs between countries. In Germany, in the domain of mathematics, 17.5% reached this competence level in PISA 2012. This number declined to 13.0% in 2015 and stayed there in 2018 (OECD, 2014, 2016, 2019b).

In contrast to these single cut-off scores on one specific achievement indicator, the consideration of different indicators and their respective cut-off scores gained importance for the identification of high-achieving student subgroups, as the multidimensional perspective on high-achieving students gained popularity. In the cut-off score approach, this comes along with three leading possibilities of combination rules. First, the conjunctive model, also known as the “and-approach”, demands the minimum requirement for high-achievement in each used identification indicator. Second, the disjunctive or complementary model, also known as the “or-approach”, demands the minimum requirement for high-achievement in at least one used identification indicator. Third, in the compensatory model, also known as the “mean-approach”, a compensation between the used identification indicators is possible, and the mean score across these indicators has to fulfill the minimum requirement. The decision to use one of these models largely influences the number of students who are identified as high-achieving. For the conjunctive model, the identified number of students strongly decreases with the increase of the number of indicators used and/or the decrease of correlations between the indicators. The opposite is true for the disjunctive model. Overall, the disjunctive model leads to a more diverse student sample, including a higher range of test scores than the conjunctive model (McBee et al., 2014). Two main reasons led to the use of the disjunctive model in the current research project in order to identify high-achieving student subgroups. First, the investigation of a broad sample of high-achieving and potentially high-achieving students, which might show, in addition to level differences, qualitative differences in the used identification indicators and can help to solve societal problems (KMK, 2016a; OECD, 2020) was of interest. Second, the interest lay in promoting high-achieving students in the mixed-ability classroom using a more inclusive perspective that is helpful for more students than only the highly gifted ones (Dixson et al., 2020). As a false positive identification of high-achieving students has hardly any negative consequences, when they are promoted through

DI inside mixed-ability classrooms, the identification of a broad high-achieving student sample is more advantageous compared to a too strict one. In this case, the disjunctive model is the most appropriate choice (McBee et al., 2014).

Despite the frequent use of cut-off scores and its easy implementation, this methodological approach was criticized for various reasons. On the one hand, methodological problems, such as loss of effect size or spurious significances can occur. Therefore, variable-centered approaches, such as regression and correlation analysis with non-dichotomized variables, were recommended (MacCallum et al., 2002). On the other hand, one criticism regarding the content is that information about individual differences is lost when a continuous variable is split into two parts. Imagine dichotomizing the IQ scale into one student group scoring 130 or higher and another scoring below 130. Before the dichotomization, we can state that a student with an IQ of 130 is more similar to a student with an IQ of 135 compared to a student with an IQ of 140. After the dichotomization, students with an IQ of 130, an IQ of 135, and an IQ of 140 are considered equal (MacCallum et al., 2002). As a consequence, these students would receive the same promotion at the same level. Further, the decision which cut-off score is used has often been criticized as arbitrary (Hickendorff et al., 2018; Stoeger, 2009). One way of overcoming the criticism of arbitrary cut-off scores, while maintaining the advantages of a person-centered over a variable-centered approach, are latent profile analyses (LPAs).

2.2.2 Latent Profile Analyses

More recently, person-centered approaches have gained importance for studying individual differences between and within individuals. Variable-centered approaches focus on the (linear) relationship between variables and how these variables change (e.g., the relationship between achievement scores and motivation). A variable-centered approach, which assumes linear relationships, has significant limitations as it ignores interaction effects and nonlinear relationships and thus also the unlikelihood that human behavior is mainly linear (Bogat et al., 2016; Hickendorff et al., 2018).

In comparison, person-centered approaches focus on describing subgroups of individuals and how these are related to other variables (Bogat et al., 2016; Hickendorff et al., 2018; Howard & Hoffman, 2018; Mammadov et al., 2016). Hence, if the focus lies on subgroups of individuals instead of a universal rule on an aggregated level, the use of a person-centered approach might be more appropriate (Bogat et al., 2016). One often used person-centered approach are LPAs that categorize individuals into homogeneous subgroups. This categorization is based on individuals' response patterns across different continuous and categorical variables and aims to maximize the homogeneity within, and the heterogeneity between, subgroups (Howard & Hoffman, 2018; Lazarsfeld & Henry, 1968; Mammadov et al., 2016).

For instance, it could be shown that on an aggregated level, a high correlation (r = .82) between two operationalizations of mathematical competence in PISA 2012 and the National Assessment Study (NAS) 2012 were found. However, on an individual level, large intraindividual discrepancies between those two conceptualizations partly existed. Students with a PISA test score of 500 had test scores between approximately 380 and 610 in NAS (Ehmke et al., 2017). Extreme discrepancies within achievement profiles are especially relevant for high-achieving students (Lohman et al., 2008). Hence, person-centered approaches that can detect high-achieving students' intraindividual differences across different achievement indicators are crucial. Despite the advantages of LPAs, they are also criticized. One criticism is the exploratory character of LPAs and that the detected subgroups could also result from methodological issues (Grimm et al., 2017).

In order to empirically investigate how high-achieving student subgroups differ, depending on the methodological approach used (i.e., cut-off score vs. LPA), the more traditional cut-off score approach (e.g., Carman, 2013; Neuendorf et al., 2022) and the more current development in the field of giftedness research, the LPA approach (Mammadov et al., 2016), was used with the same high-achieving student sample and the respective classification results were compared. A comparison of the two approaches provides important information from a methodological perspective on whether both approaches produce comparable results despite their respective criticisms and, if not, where differences in the results between the two approaches can be found. Advantages and disadvantages of both approaches in different educational settings are discussed.

2.3 Summary

This chapter summarizes theories, models and empirical research regarding the classification of high-achieving students in an educational context. One important topic in this context is the multidimensional understanding of high-achieving students, including domain-general and domain-specific achievement, as well as motivational-affective characteristics. This multidimensional perspective and the lack of consensus regarding the classification of high-achieving students (Neuendorf et al., 2022) is associated with different challenges that need to be resolved. One important open question relates to the comparability of the different identification indicators used and to what extent high-achieving students differ across these indicators. Instead of considering high-achieving students as one homogeneous group, different methodological approaches are used to identify high-achieving student subgroups. Here, the cut-off score and the LPA approach play an important role. Both approaches have advantages and disadvantages in the educational research and practice. However, it still remains an open question if both approaches show comparable classification results when the aim is to identify high-achieving student subgroups.

Further, the multidimensional understanding of high-achieving students results in a differentiated view regarding their strengths and weaknesses. However, the question of how these differences can be taken into account in educational practice is still discussed (Barbier et al., 2022; García-Martínez et al., 2021). Based on the knowledge of the heterogeneity between high-achieving students and, therefore, the importance of individualized instruction, the following chapter discusses potential and limitations of DI for high-achieving students in mixed-ability classrooms.

3 Promotion of High-Achieving Students

The identification of high-achieving students' specific characteristics and needs has only added value for education if schools actually adapt their teaching to it and offer students individualized in-school support (Dixson et al., 2020). Hence, this chapter first briefly describes the changing role of high-achieving students in the educational context, followed by a short introduction to the three main forms of support measures for high-achieving students, namely enrichment, acceleration, and grouping (García-Martínez et al., 2021; Weigand, 2019). The focus of this chapter lies on the importance of supporting high-achieving students in mixed-ability educational settings through DI to meet their varying needs.

3.1 The Role of High-Achieving Students in the Educational Context

One dilemma teachers often face in the educational context due to limited resources is to decide whose needs should be prioritized and how resources should be distributed (Tirri & Laine, 2017). According to Deutsch (1985), there are three forms of distributive justice in co-operative relations. If the focus is on economic productivity, equity is the leading principle. If the focus is maintaining positive social relations, the leading principle is equality and if the focus is on personal development and welfare, the leading principle is personal need. Looking at past political developments and research, the focus has often been on struggling students (e.g., Fischer & Müller, 2014; KMK, 2010; Simpson et al., 2004; Tirri & Laine, 2017), despite the right of all students to get appropriate education according to their individual potential (Borland, 2005; KMK, 2009). As a consequence, the promotion of potentially high-achieving and already high-achieving students has been neglected (García-Martínez et al., 2021; Kunze et al., 2019; Moltzen, 2011). One reason, therefore, might have been the misconception that high-achieving students do not need extra promotion as they will do it on their own (Moltzen, 2011). This misconception appeals to school staff and politicians as they can focus on struggling students and also do not have to address high-achieving students' needs and share scarce resources (Moon, 2009). As struggling students' needs are believed to be more urgent, they are prioritized and, therefore, the needs of high-achieving students are not met or are only done so to a lesser extent (Tirri & Laine, 2017). Further, high-achieving students might appear appropriately fostered in an underchallenging learning environment, as they might still achieve above average performance. However, they run the risk of developing "maladaptive motivational beliefs that will sabotage their resilience when they encounter more challenging coursework in the future" (Moon, 2009, p. 274).

Although, especially in Germany, the promotion of high-achieving students gained importance relatively late (Kunze et al., 2019; Stumpf, 2012), the neglect of high-achieving students has changed. Political resolutions aiming for the identification and promotion of high-achieving students were resolved (KMK, 2015, 2016a; Kunze et al., 2019). These resolutions resulted in the large ten-year project "Leistung macht Schule" (LemaS; Promoting Excellence in School Education) with schools across all German school types and 22 subprojects in two main areas. Main area one focuses on the development of school mission statements that include joint values and aims regarding a potential-oriented school culture and the development of cooperative networks. The second main area is strongly related to this dissertation. The focus lies on the development of diagnosis-based individualized promotion of high-achieving and potentially high-achieving students in regular mixed-ability classrooms in order to better develop their full potential. A scientific evaluation of the project is also taking place, wherefore important empirical findings for the promotion of high-achieving students in Germany can be expected in the next few years (Weigand et al., 2020).

This change is important because of two main reasons at an individual and a societal level. On the one hand, educational promotion approaches for all students also include high-achieving students (KMK, 2009) as all students have the right to receive appropriate education according to their needs (Borland, 2005). Hence, the claim that "professional ethics in teaching calls for equal attention to all students and equitable educational opportunities and experiences" occurs (Tirri & Laine, 2017, p. 250). On the other hand, in order to improve the performance of an education system and therefore realize the potential of a society, high-achieving students play an important role (OECD, 2019b). However, even though the view that high-achieving students need to be educated according to their individual needs has gained support and attention (e.g., Moltzen, 2011), the question of how the concrete implementation of this support should look is still under discussion. Diverse educational interventions have been used so far (Barbier et al., 2022; Dai et al., 2011; García-Martínez et al., 2021). The focus thereby should be on individualized support according to the characteristics and needs of high-achieving students in order to promote their educational outcomes as best as possible (García-Martínez et al., 2021).

3.2 Enrichment, Acceleration, and Grouping

The three international and national most widespread forms of support measures for high-achieving students in the educational context are called enrichment, acceleration, and grouping (Heller, 2001; Walsh et al., 2012; Weigand, 2019). Figure 1 gives an overview of different exemplary support measures in these three areas and points out that these measures can take place inside and outside of mixed-ability classrooms. The focus of this dissertation is on support measures inside mixed-ability classrooms that can be part of all three areas (grey frame).

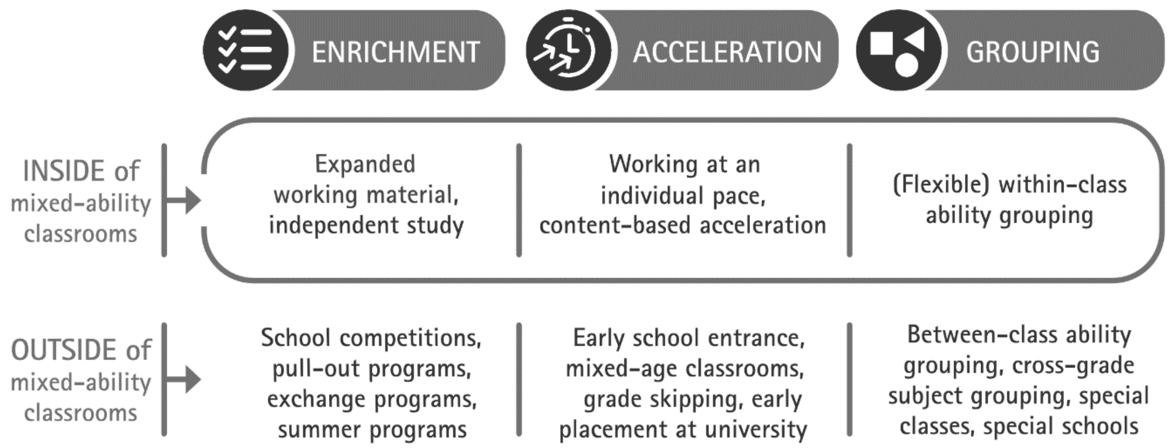


Figure 1. Support Measures for High-Achieving Students Inside and Outside Mixed-Ability Classrooms (translated and adapted from Ziernwald et al., 2020 based on KMK, 2009, 2015).

Enrichment focuses on the enlargement or/and deepening of learning, for example, through school competitions (KMK, 2015; Stumpf, 2012). Regarding enrichment, a meta-analysis has shown, on average, large positive effects on academic achievement and medium positive effects on socioemotional development. The effect sizes varied depending on the type of enrichment program and students' age. For example, older students benefited more than younger students (Kim, 2016).

Acceleration, in contrast, describes a faster progression through the school career, for example, through grade skipping or early entrance at school (KMK, 2015; Plucker & Callahan, 2020; Stumpf, 2012). Research syntheses found (small) positive effects of different acceleration approaches on students' academic achievement and their motivational-affective characteristics (Steenbergen-Hu et al., 2016; Steenbergen-Hu & Moon, 2011).

The third area, grouping, focuses on the aggregation of students, for example, based on their ability to create more homogeneous subgroups (KMK, 2015). This support measure has been studied in great detail for about 100 years. A second-order meta-analysis summarized the findings of 13 ability-grouping meta-analyses. It concluded that within-class grouping, cross-grade subject grouping, and special grouping for the high-achieving students showed small positive effects on students' academic achievement, whereas between-class grouping showed no effects. These effects were found independent of students' ability levels (Steenbergen-Hu et al., 2016). However, researchers repeatedly stated that grouping alone is not enough. In order to unfold the positive effects of grouping, grouping should be accompanied by appropriately adapted instruction for the different student groups (Kulik & Kulik, 1992; Plucker & Callahan, 2020; Tieso, 2003).

Although this triad exists, in practice, many support measures for high-achieving students combine different facets and are therefore not clearly distinguishable from each other (Heller, 2001; Stumpf, 2012). Next to these three areas of support measures, integrated promotion inside mixed-ability classrooms is discussed for high-achieving students. One key element thereof is DI (KMK, 2015). However, research to date regarding educational practices within mixed-ability classrooms for high-achieving students is still sparse (Barbier et al., 2022). In the following, a short introduction into external and internal differentiation is given, along with an explanation of the focus on DI as part of internal differentiation inside mixed-ability classrooms.

3.3 Differentiated Instruction in Mixed-Ability Classrooms

The term differentiation is often divided into external and internal differentiation (Bönsch, 2004). On the one hand, an important construct describing external differentiation is tracking, that can take place between schools through different school types and within schools through different streams. Their common aim is to form relatively homogeneous fixed subgroups of students based on their prior achievement in order to have students with similar prior achievement grouped within school types and school streams. In Germany, this tracking system is one of the most strict ones (Retelsdorf et al., 2012; Trautwein et al., 2006). A preference for teaching to the middle through teacher-centered teaching leads to the attitude that heterogeneity is not seen as gain but as a problem that should be removed (Trautmann &

Wischer, 2011). On the other hand, internal differentiation focuses on flexible student subgroups based on their achievement differences within regular mixed-ability classrooms (Dumont, 2018). One integral part of internal differentiation is DI during the lesson and can be defined as “the adaptation of content, process, product, learning environment or learning time based on information about students’ readiness or another relevant student characteristic (such as learning preference or interest) to better address students’ learning needs” (Smale-Jacobse et al., 2019, p. 3). This adaptation seems especially important for high-achieving students as 73% of the teachers in a study stated that high-achieving students are too often bored and demanded too little at school and do not have sufficient possibility to grow (Loveless, 2008). This teacher perspective is in accordance with the high-achieving students’ perspective. Only 14.2% perceived their mathematics lessons, and 15.3% their science lessons, as highly cognitive activating (Rieser et al., 2016). Hence, despite external differentiation, DI is necessary due to several reasons. First, the tracking in Germany, for example, only exists for secondary education. Therefore, in the untracked primary education system, a large heterogeneity exists by nature. Second, in secondary education, despite tracking, a large achievement heterogeneity exists (Bohl et al., 2012). For example, in PISA, for 9th graders at the highest German school track, the reading and mathematics achievement of the middle 90% of students spans across four to five proficiency levels (Reinhold et al., 2019; Weis et al., 2019). Third, all over the world, inclusive education gained importance (Armstrong, 2011; Lindner & Schwab, 2020). The UNESCO program strives for inclusive education for all people (General Assembly of the United Nations, 2019; UNESCO, 2017). The favor for inclusive education also follows researchers who stated that there is only little support from research for this inflexible grouping through tracking (Plucker & Callahan, 2020) or even found (partially) unwanted side effects of tracking (e.g., Oakes, 2005; Retelsdorf et al., 2012). In line with this claim for inclusive education, a growing heterogeneity of students in the regular mixed-ability classrooms can be observed (Civitillo et al., 2016; Fischer, 2019; Subban, 2006). Hence, the one-size-fits-all approach appears inappropriate and a shift towards DI can be observed (Lindner & Schwab, 2020).

Further, the idea of promoting high-achieving students inside mixed-ability classrooms through DI involves various advantages. First of all, most high-achieving students are educated most of the time in mixed ability classrooms (Dreeszen, 2009; White et al., 2003; Yuen et al., 2018) and can therefore be challenged appropriately on a daily basis and not only during a short period of time, as is often the case with pull-out programs (Dreeszen, 2009; Gubbins, 2013). Second, high-achieving students themselves, parents, as well as teachers supported the promotion in inclusive educational settings, whereas segregated promotion approaches such as special classes only received little support (Shayshon et al., 2014; Sparfeldt et al., 2004). Third, including high-achieving students in mixed-ability classrooms can have positive effects for themselves as well as their not-as-high-achieving identified classmates (Ninkov, 2020). Last but not least, a special advantage of the inclusive education of high-achieving students in mixed-ability classrooms is that the identification problem of high-achieving students is no longer of such great importance. In the case of DI, the identification of high-achieving students does not result in fixed long lasting high-stakes decisions (e.g., Acar et al., 2016) with strong detrimental effects for misidentified high-achieving students. The assignment of students to certain DI approaches is quite flexible compared to external differentiation measures, where assignment to special classes, for example, is quite fixed (Dumont, 2018).

Further, DI does not foster high-achieving students *or* not as high-achieving identified students *or* low-achieving students, but it could be beneficial for *all* students in mixed-ability classrooms (Borland, 2005; Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019). Hence, high-achieving minority students (e.g., economically disadvantaged, ethnic minorities) who are less often labeled as high-achieving and therefore less often considered for special out-of-school promotion programs can also be supported within inclusive educational settings (Maker, 1996; Moltzen, 2011).

Regarding the effectiveness of DI, three current research syntheses and a literature review found positive effects on students' achievement and emphasized the potential of DI in order to fulfill students individual needs (Bondie et al., 2019; Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019). However, as we know from the aptitude-treatment interaction research, positive effects found in the whole sample do not automatically hold true for a subsample (e.g., high-achieving students; Connor et al., 2011; Cronbach & Snow, 1977). A systematic review on educational practices for high-achieving students found that DI positively influenced their cognitive outcomes. However, research support for DI was less clear compared to other educational practices, such as dynamic feedback and self-regulated learning (Barbier et al., 2022). Accordingly, a systematic and comprehensive overview of research regarding the specific potential of DI for high-achieving students in mixed-ability classrooms is needed.

Further, although DI is a key aspect of effective instruction and teachers in general appreciate DI's potential for teaching, they also perceive difficulties while implementing it (Kahmann et al., 2022). In accordance with these perceived difficulties, we know from pivotal research from the 1990s that DI is not used on a regular basis for high-achieving students (Archambault et al., 1993; Westberg et al., 1993). This result was again confirmed at least for America in a more recent literature review (Reis & Renzulli, 2010). Hence, if DI is not regularly implemented for high-achieving students, it cannot unfold any existing potential. One factor that influences the (intended) use of instructional approaches are practitioners' beliefs and expectations, such as the importance or usefulness of a certain approach (Bondie et al., 2019; Gebauer & McElvany, 2017; Trivette et al., 2012). Therefore, to create a broad picture of DI for high-achieving students that goes beyond the impact of DI on high-achieving students' outcomes, the frequency of use, the perceived utility, as well as barriers and facilitators for the successful implementation of DI should also be systematically investigated.

3.4 Summary

This chapter first described why high-achieving students were a neglected group in the educational context (e.g., misconceptions) for so long and have now moved more into focus (e.g., right to appropriate education for all students). Next, an overview of different support measures that can be used for high-achieving students in educational contexts was given. Thereby, the focus was on DI in mixed-ability classrooms. On the one hand, inclusive education generally gained importance and now also includes high-achieving next to struggling students (García-Martínez et al., 2021; Lindner & Schwab, 2020; Tirri & Laine, 2017) and, on the other hand, promotion of high-achieving students through DI in mixed-ability classrooms can have several advantages (e.g., regular daily promotion). So far, many different theoretical models as well as teaching examples regarding the education of high-achieving students exist (e.g., Kaplan, 1986; Maker & Schiever, 2010). However, the overall empirical evidence regarding the promotion of high-achieving students and especially research regarding support measures inside mixed-ability classrooms is still sparse (Barbier et al., 2022; García-Martínez et al., 2021). It was argued that a systematic overview regarding the impact of DI on high-achieving students' achievement and motivational-affective outcomes is necessary as the treatment-aptitude interaction shows that results that are valid for the overall population do not automatically hold true for a certain subgroup (Connor et al., 2011; Cronbach & Snow, 1977). Further, it is important to investigate the open question of which hindering and facilitating conditions for the successful implementation of DI exist (e.g., perceived utility) as effectiveness alone will not automatically cause successful implementation of DI in a lesson (Missett et al., 2014).

4 The Present Research

Two main research fields of giftedness research were the starting point of the current research project: (1) classification of high-achieving students, including identification and characterization of high-achieving student subgroups and (2) individualized in-school support with a focus on promotion of high-achieving students within an inclusive educational setting through DI. Figure 2 summarizes the framework of the current dissertation and shows the added value in the giftedness research field of classification as well as individualized in-school support.

These aspects are highly important from an individual and a societal perspective and are further underlined through a political resolution stating that dealing with heterogeneity is one crucial topic in the upcoming years in educational research and practice, which includes individualized support in heterogeneous learning groups focusing on struggling as well as high-achieving students (KMK, 2016b). Thereby, a prerequisite for the appropriate promotion of high-achieving student subgroups, including the use of individualized strategies, is knowledge about the classification of high-achieving students and their characteristics and needs. As high-achieving students are not one homogeneous student group (Moltzen, 2011; Reis & Renzulli, 2009), a detailed description of the heterogeneity of different subgroups of high-achieving students is needed. The knowledge about the characteristics of high-achieving student subgroups supports "the selection of the most appropriate educational interventions for the characteristics of each type of student" (Castejón et al., 2016, p. 167). In order to develop this knowledge about the heterogeneity of high-achieving students and to provide individualized promotion according to their needs, a multidimensional perspective on high-achieving students is essential (e.g., Heller et al., 2005; Stoeger, 2009). Further, the description of high-achieving students' heterogeneity and needs has only added value for the educational context if teachers actually adapt their instruction to their needs and offer individualized in-school support (Dixson et al., 2020).

For the investigation of this dissertation's framework (Figure 2), different theoretical strands from intelligence research (e.g., Cattell, 1943), giftedness research (e.g., Heller et al., 2005; Stoeger, 2009), and educational research (i.e., Neidorf et al., 2006; Saß et al., 2017) are brought together. These theoretical perspectives are complemented with important methodological perspectives that are, on the one hand, more traditional (i.e., cut-off score approach; Carman, 2013; Neuendorf et al., 2022) and, on the other hand, more innovative (i.e., LPA approach; Mammadov et al., 2016) in the field of high-achieving students. Last but not least, the practical perspective is particularly emphasized through the investigation of DI as a promising support measure (Barbier et al., 2022; Deunk et al., 2018; Smale-Jacobse et al., 2019) for high-achieving students in mixed-ability classrooms. Thereby, different aspects with high practical relevance (i.e., impact on high-achieving students' outcomes, frequency of use, perceived usefulness, and barriers and facilitators for successful implementation) were considered. Accordingly, the superordinate research question (RQ_0) is the following:

RQ_0: How can the multidimensional characteristics of high-achieving students be classified and appropriately taken into account through differentiated promotion in inclusive educational settings?

In order to gain new insights and to close existing research gaps in the classification of high-achieving students, a multidimensional perspective (i.e., domain-general achievement, domain-specific achievement, and motivational-affective characteristics) and two methodological approaches were used (Paper A: Ziernwald, Schiepe-Tiska, & Reiss, 2022). The main aim was to describe the existing heterogeneity between high-achieving students through the identification of high-achieving student subgroups and to describe their varying characteristics in a second step. For the identification of the high-achieving student subgroups, the cut-off score and the LPA approach were used in order to see how these two methodological approaches agree in their findings and can thus be generalized. Accordingly, the following two research questions were investigated: (1) Do both methodological approaches (cut-off score and LPA approach) derive comparable high-achieving student subgroups (RQ_{1a}) and (2) How can these subgroups be characterized regarding their motivational-affective characteristics (RQ_{1b})?

To gain new insights and close existing research gaps regarding individualized in-school support measures, the promotion of high-achieving students in mixed-ability classrooms through DI was investigated (Paper B: Ziernwald, Hillmayr, & Holzberger, 2022). As DI aims to consider students' differences while building on their commonalities (Tomlinson, 2001), this educational approach might be useful in an inclusive educational context in order to develop the full potential of all students, also including the high-achieving ones. To see, on the one hand, what is already known from research and which research gaps still exist regarding DI for high-achieving students and, on the other hand, to derive implications for research and practice, a comprehensive, systematic mixed-methods review was conducted. Therefore, the impact of DI on high-achieving students' achievement and motivational-affective characteristics as well as conditions regarding its implementation were investigated. Thus, four research questions regarding (1) the impact of DI on high-achieving students' outcomes (RQ_{2a}), (2) the extent of DI use (RQ_{2b}), (3) the perceived usefulness of DI from teachers' and high-achieving students' perspective (RQ_{2c}), as well as (4) barriers and facilitators while implementing DI for high-achieving students (RQ_{2d}) were addressed.

A more detailed description of the two studies can be found in chapter 5 and the journal articles (Appendix A and B). Next to the current dissertation's two leading publications, a practice-oriented brochure was published for teachers, politicians, and the interested public. Here, an overview of support measures of high-achieving students, including research findings as well as information portals and training opportunities in the field of gifted education, were presented (Ziernwald et al., 2020). Through this practice-oriented brochure and different exchange formats with teachers and teacher educators, the research results obtained should also be taken into account and applied in educational practice.

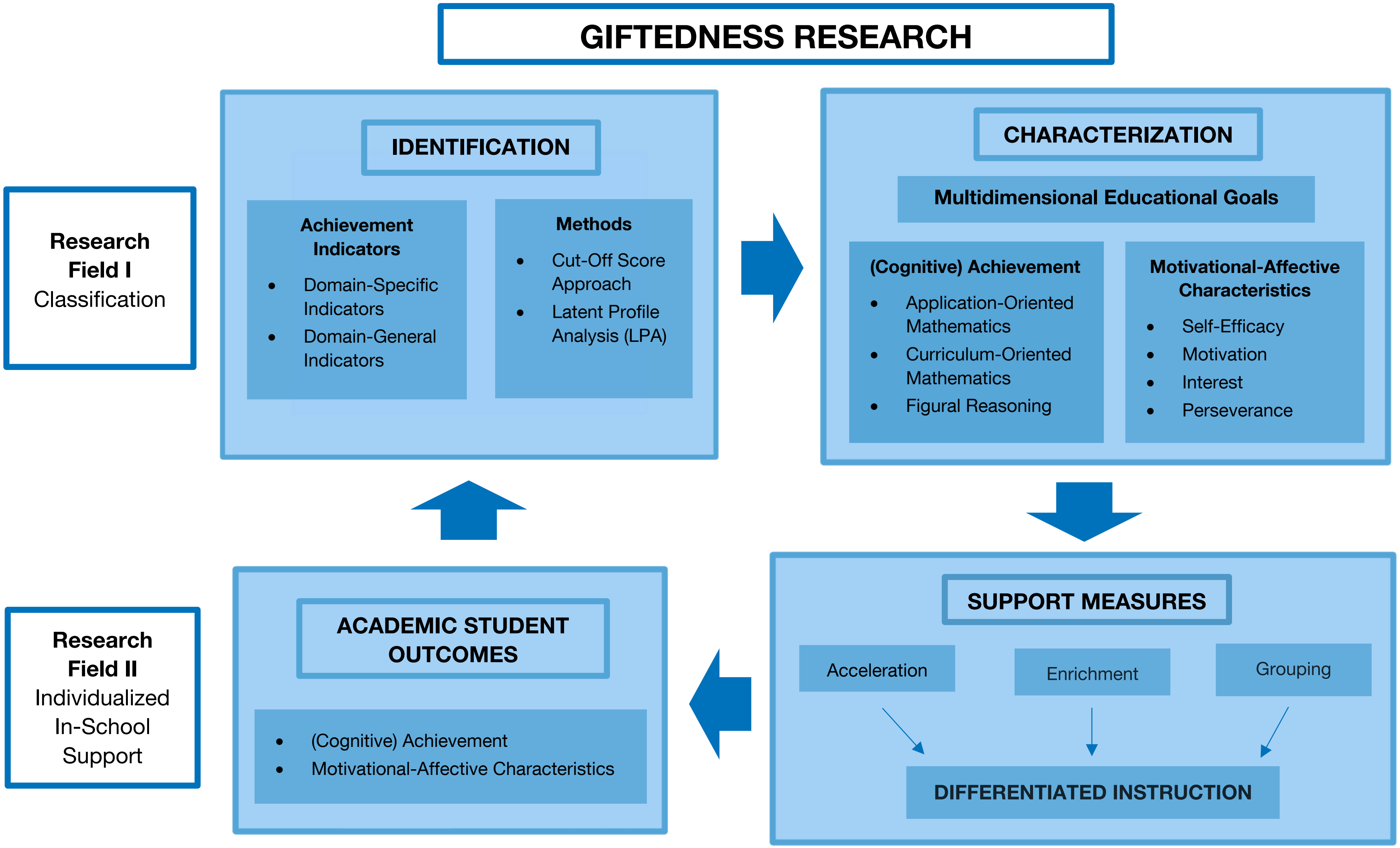


Figure 2. Framework of the Current Dissertation.

5 Summary of Publications

This dissertation contains two publications contributing to giftedness research from different perspectives and with different methodological approaches. Paper A aimed to identify high-achieving student subgroups and investigate their heterogeneity from a multidimensional perspective. The paper was published in the international, peer-reviewed journal *Learning and Individual Differences*. The dissertation's author is the first and corresponding author of the publication and was primarily responsible for the conceptualization, methodology, analysis, writing, and review of the publication (75%). The co-authors, Professor Dr. Anja Schiepe-Tiska (20%) and Professor Dr. Kristina Reiss (5%) supported the different steps through discussion and added remarks and additions to the different sections of the paper.

Paper B aimed to summarize and discuss existing research and reveal research gaps regarding DI for high-achieving students, wherefore a systematic mixed-methods review was conducted. The paper was published in the international, peer-reviewed journal *Journal of Advanced Academics*. The dissertation's author is the first and corresponding author of the publication and was primarily responsible for the conceptualization, methodology, analysis, writing, and review of the publication (70%). The co-authors, Delia Hillmayr (10%) and Professor Dr. Doris Holzberger (10%) supported the different steps through discussion and added remarks and additions to the different sections of the paper. Delia Hillmayr and others (10%) further supported the screening and double-coding of the (included) primary studies.

A summary of the two studies is provided below. A detailed description of the two studies is presented in the corresponding journal articles (see Appendix A and B). The discussion of the main findings, methodological strengths and limitations, as well as practical implications of the studies, can be found in chapter 6.

5.1 Identification and Characterization of High-Achieving Student Subgroups (Paper A)

Ziernwald, L., Schiepe-Tiska, A., & Reiss, K. M. (2022). Identification and characterization of high-achieving student subgroups using two methodological approaches: The role of different achievement indicators and motivational-affective characteristics. *Learning and Individual Differences*, *100*, Article 102212. https://doi.org/10.1016/j.lindif.2022.102212

5.1.1 Research Aims

Paper A takes as its starting point the paucity of research on how to identify and characterize high-achieving student subgroups and the extent to which different methodological classification approaches (i.e., cut-off score approach, LPA approach) yield generalizable results (Acar et al., 2016; Carman, 2013; Neuendorf et al., 2022). This research gap was clarified through the description of the different identification indicators, as well as identification methods used and how this influences the ability to generalize results. Therefore, a multidimensional perspective on high-achieving students was chosen (e.g., Heller et al., 2005; Stoeger, 2009). Besides domain-general and domain-specific achievement indicators (e.g., Cattell, 1943), domain-general and domain-specific motivational-affective characteristics (e.g., Gutman & Schoon, 2013) were also considered. For the domain-specific achievement indicators, mathematics was chosen, as it offers a particularly valuable complement to domain-general achievement (i.e., fluid intelligence). On the one hand, the mastery of complex quantitative reasoning becomes increasingly important for a large number of educational and occupational purposes (Dowker et al., 2016; OECD, 2018) and can be considered as an essential condition of social participation and lifelong learning (Ehmke et al., 2020). On the other hand, if one domain-specific area was used to operationalize high-achieving students, mathematics is by far the most often chosen domain (Neuendorf et al., 2022).

The first goal of Paper A was to investigate to what extent the cut-off score and the LPA approach using three different achievement indicators (i.e., application-oriented mathematical competence, curriculum-oriented mathematical competence, and figural reasoning) result in comparable, and therefore generalizable subgroups of high-achieving students (RQ_{1a}). The second goal was to characterize the different detected high-achieving student subgroups regarding their motivational-affective characteristics (i.e., perseverance, mathematics self-efficacy, interest, and instrumental motivation) in more detail (RQ_{1b}). With these two main research questions, the paper aimed to contribute to the description of the heterogeneity of high-achieving students and to sensitize regarding the consequences of selecting certain methodological approaches and indicators to identify high-achieving student subgroups. The description of the heterogeneity between high-achieving student subgroups and how they can be characterized through achievement indicators and motivational-affective characteristics should facilitate individualized and appropriate gifted education (Castejón et al., 2016).

5.1.2 Methodology

To answer the first research question of whether the cut-off score and the LPA approach result in comparable high-achieving student subgroups, and also the second research question of which characteristics these different subgroups of high-achieving students have, a large dataset with a broad range of high-achieving students is useful. As the PISA study is "the most comprehensive and rigorous international programme to assess student performance and to collect data on the student, family and institutional factors that can help explain differences in performance" (OECD, 2019a, p. 13), this dataset was appropriate for answering our research questions. Thereby, the PISA 2012 dataset has a special advantage, as a linkage to the German NAS, which investigated curriculum-based competences and figural reasoning was possible (Pant et al., 2013).

Our final sample consisted of N = 1563 high-achieving 9th graders (45.6% females, 78.7% attending the highest German school track), with a mean age of M = 15.30 years (SD = 0.46). Data collection took place on two consecutive days. On the first day, PISA's mathematics test and the background questionnaire–including the motivational-affective characteristics–were administered, followed by the NAS' mathematics test and figural reasoning on the second day.

Figure 3 displays the three main analysis steps, beginning with detecting subgroups through the cut-off score (Step 1a) and the LPA approach (Step 1b). In Step 2, we used 50 imputed data sets to analyze the high-achieving student subgroups' motivational-affective characteristics with regression analyses.

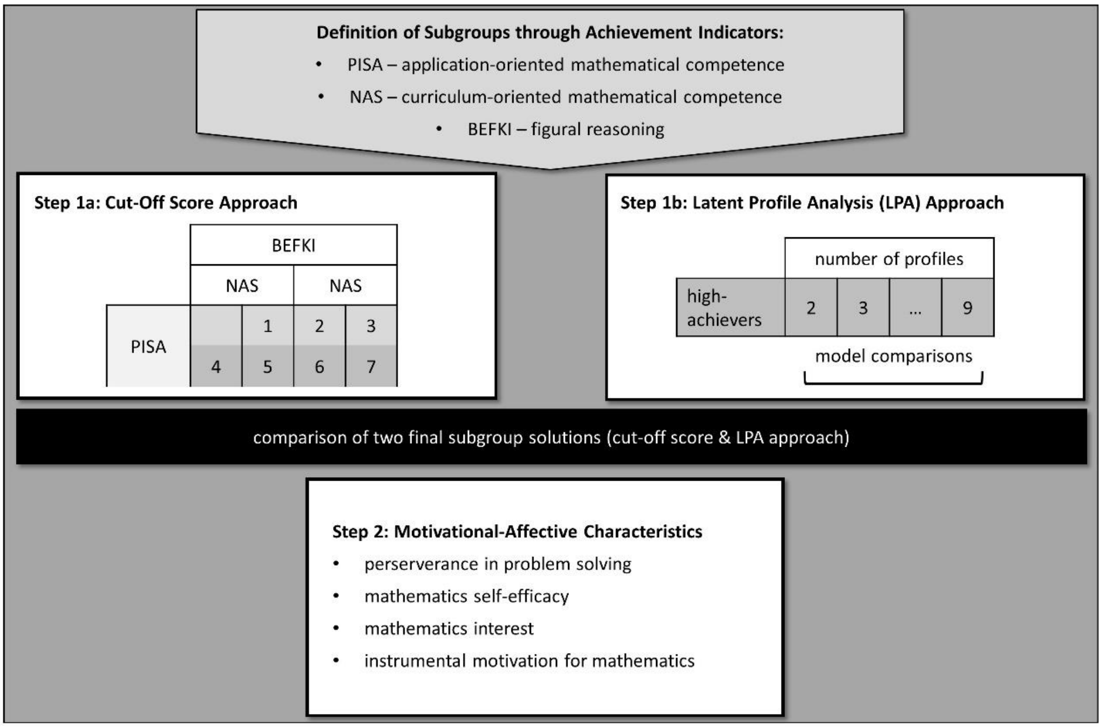


Figure 3. Analysis Steps for the Identification and Characterization of High-Achieving Student Subgroups.

5.1.3 Main Results

Valuable empirical insights from a theoretical, methodological and practical perspective were gained. The direct comparison of the two methodological approaches showed that, at an aggregated level, the four largest subgroups of high-achieving students (i.e., Top Performers, Overachievers, Underachievers, Operators) were identified through both approaches. However, at an individual level, the overlap is 59%. This medium overlap was found regardless of whether the 5-profile or 7-profile solution was used in the LPA approach and leads to the assumption that the cut-off score and the LPA approach cannot be used interchangeably for a considerable part of high-achieving students at the individual level.

From a content perspective, the large heterogeneity within the sample of high-achieving students is remarkable. All three achievement indicators showed a large range of values, and different relationships between the achievement indicators across high-achieving student subgroups were observed. This heterogeneity in the achievement indicators is also reflected through the motivational-affective characteristics, which supports the assumption that high-achieving students are *not* a homogeneous group and a multidimensional perspective is necessary and beneficial.

5.2 Promotion of High-Achieving Students through Differentiated Instruction (Paper B)

Ziernwald, L., Hillmayr, D., & Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms–A systematic review. *Journal of Advanced Academics*, *33*(4), 540–573. https://doi.org/10.1177/1932202X221112931

5.2.1 Research Aims

Paper B systematically and comprehensively summarized, analyzed, and discussed the existing research regarding DI for high-achieving students in mixed-ability classrooms. The starting points were various theoretical contributions regarding the usefulness of DI (e.g., Kaplan, 1986; Maker & Schiever, 2010; Tomlinson, 2001), empirical research syntheses emphasizing the potential of DI for students in general (e.g., Bondie et al., 2019; Deunk et al., 2018; Smale-Jacobse et al., 2019), and a diverse set of primary studies, investigating DI for high-achieving students from different perspectives.

The main goal of the systematic mixed-methods review was to gain a comprehensive overview regarding the impact of DI on high-achieving students' educational outcomes (RQ_{2a}) and the conditions for successful DI implementation for high-achieving students in mixed-ability classrooms (RQ_{2b}, RQ_{2c}, RQ_{2d}). Accordingly, empirical knowledge regarding the impact of DI on high-achieving students' achievement and motivational-affective outcomes as well as implementation conditions of DI, were contributed.

5.2.2 Methodology

Research syntheses are an appropriate method for gaining a comprehensive overview of a research field and for detecting possible research gaps (Beelmann, 2014). Figure 4 displays the four main analysis steps of the conducted systematic mixed-methods review. The first step was a comprehensive literature search, including database searches, a hand search, and a snowball search resulting in overall *N* = 7736 hits. For the second step, we screened the title and the abstract of each detected study from the database and hand search and included or excluded the study based on predefined inclusion and exclusion criteria. Studies which passed step two were checked again in step three with the full text regarding their eligibility. If a study was eligible, the relevant information to answer the research questions was extracted. In the fourth step, we calculated effect sizes and conducted thematic syntheses for the different research questions with *N* = 49 primary studies. Fifteen studies contributed to RQ_{2a}, 12 studies contributed to RQ_{2b}, 13 studies contributed to RQ_{2c}, and 22 studies contributed to RQ_{2d}.

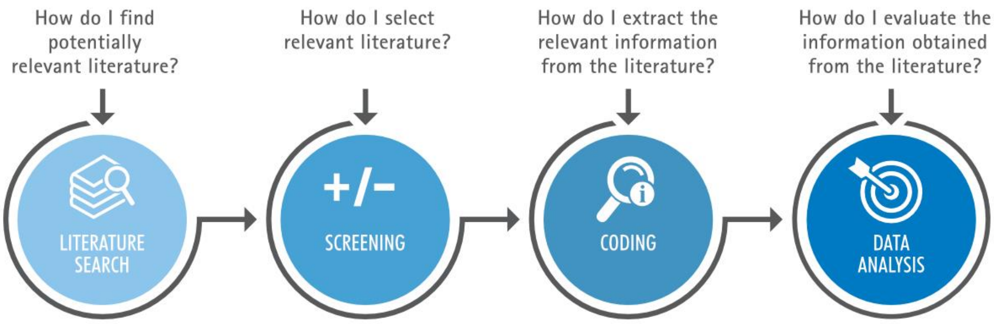


Figure 4. Main Steps of a Systematic Review.

As the aim was to gain the most comprehensive overview of DI for high-achieving students, and as the four different research questions required different perspectives (e.g., objective test results vs. subjective opinions), a systematic mixed-methods review was conducted. This type of review includes qualitative, quantitative, and mixed-methods primary studies and therefore allows the use of contributions and advantages of the qualitative and the quantitative methods. Thus, complex research questions can be answered in more detail and more completely (Heyvaert et al., 2013).

5.2.3 Main Results

Regarding the impact of DI on high-achieving students, it can be stated that DI is a promising approach for education in mixed-ability classrooms regarding achievement and motivational-affective characteristics. Further, it was found that DI is not used proactively and regularly for high-achieving students in mixed-ability classrooms. However, teachers and high-achieving students themselves rated DI predominantly as useful and appropriate for the promotion of high-achieving students. This discrepancy between the relatively low frequency of usage and the high subjective usefulness can possibly be explained by the various barriers (e.g., lack of resources, misconceptions, focus on struggling students) the school staff stated when implementing DI for high-achieving students in mixed-ability classrooms. In addition to these barriers, facilitators for the implementation of DI in mixed-ability classrooms were detected. Leadership support and professional staff development, for example, might help to overcome the mentioned barriers and better realize DI's potential for high-achieving students in mixed-ability classrooms.

6 Discussion

This research project investigated the classification and individualized in-school promotion of high-achieving students innovatively and comprehensively. A multidimensional perspective on high-achieving students with different achievement indicators as well as motivational-affective characteristics was used to describe high-achieving students' heterogeneity. In contrast to previous studies, which used either a cut-off score or an LPA approach, these two approaches were compared based on one sample. Further, through the systematic mixed-methods review, the research field of DI and high-achieving students was structured comprehensively. On the one hand, evidence regarding (1) the mainly positive impact on high-achieving students' outcomes, (2) the expandable frequency of use, (3) the high perceived utility, and (4) barriers (e.g., teachers' misconceptions) as well as facilitators (e.g., professional staff development) for its implementation, was found. On the other hand, research gaps (e.g., specific features of DI approaches), where future primary studies can start to develop high-achieving students' education further were revealed. Below, the main findings are discussed and placed within the current state of giftedness research. Further, strengths and limitations of the current research project as well as future directions for research and practical implications for gifted education are outlined.

6.1 Discussion of the Main Findings

The current research project contributed to the superordinate research question of how the multidimensional characteristics of high-achieving students can be classified and appropriately taken into account through differentiated promotion in inclusive educational settings. First, the two methodological approaches, namely the cut-off score and the LPA approach, were compared in order to classify high-achieving students into subgroups and describe their multidimensional characteristics. It was shown that both approaches detected on an aggregated level the four largest–and therefore for educational practice most relevant–high-achieving student subgroups from the cut-off score approach, namely Top Performers, Overachievers, Underachievers, and Operators. The first three subgroups were already known from previous research (e.g., Agaliotis & Kalyva, 2019; Hofer & Stern, 2016; Leikin et al., 2017). The smaller subgroups from the cut-off score approach (i.e., Learners, Wise Learners, and Wise Operators) were almost completely merged into the LPA Overachiever subgroup. At an individual level, the relationship between the two methodological approaches can be rated as quite high (Carmer's V = 0.67; Cohen, 1977). Nevertheless, 41% of the students were assigned to different subgroups at an individual level.

As the classification of high-achieving students and the assignment into certain subgroups can have important consequences (Acar et al., 2016), the influence of the method should always be kept in mind, especially when making decisions at the individual level. Both methodological approaches showed advantages and might be more appropriate depending on the application context. The cut-off score approach identified more qualitatively different high-achieving student subgroups and resulted in subgroups where each individual represents the overall achievement pattern of the respective subgroup. Therefore, if high-achieving students should be distributed to promotion programs that focus on different achievement areas (e.g., mathematics vs. general reasoning), the cut-off score approach might produce more appropriate subgroups. The LPA approach identified, in addition to qualitatively different high-achieving student subgroups, subgroups that differed regarding their level (Underachiever 1 vs. Underachiever 2). Thereby, subgroups were created to be as homogeneous as possible across the three achievement indicators, even if this means that not every individual represents the overall achievement pattern of the subgroup. This division into subgroups is particularly relevant in educational practice when high-achieving students should be assigned to different promotion programs that focus on different achievement levels across domains (e.g., mathematics and general reasoning program but at different levels).

In addition to these methodological and practical conclusions, this research project also provided remarkable evidence regarding the theoretical assumption that high-achieving students should be characterized through a multidimensional perspective. These results counter the misconception that high-achieving students show high achievement in all areas (Sak, 2011; Tirri & Laine, 2017). If the misconception had been true, only one subgroup, namely our Top Performer subgroup, that showed high achievement across all achievement indicators and the most positive motivational-affective characteristics would have been detected. However, this subgroup consistently contained across both methodological approaches only approximately one quarter (cut-off score: 25.40%, LPA: 23.35%) of the whole high-achieving student sample. The other subgroups showed qualitative shape and quantitative level differences from this Top Performer subgroup. The existence of these different high-achieving student subgroups and the partially large heterogeneity across subgroups within achievement indicators empirically contradict the myth that high-achieving students are one homogeneous group (Moltzen, 2011; Reis & Renzulli, 2009).

This heterogeneity in the achievement indicators was also reflected in the motivational-affective characteristics. As the Top Performers, characterized through high domain-general potential and high domain-specific achievement, showed also the most positive motivational-affective characteristics, theoretical conceptions that suggest that motivational-affective characteristics influence if potential transfers into manifest achievement (e.g., Cross & Cross, 2017; Heller et al., 2005) are empirically supported. However, as no experimental or longitudinal design was used, causal interpretations are not possible. Nevertheless, Top Performers can be considered as a target state for ideal educational promotion, as multidimensional educational goals, including achievement and motivational-affective characteristics as important student outcomes (Schiepe-Tiska, 2019), are fulfilled.

Following these strong results regarding the heterogeneity of high-achieving students regarding their achievement and motivational-affective characteristics, the question arises how these differences can be taken into account in inclusive educational settings as the one-size-fits-all approach, that is often used to date, appear inappropriate (Borland, 2005; Lindner & Schwab, 2020). As DI takes into account students differences while building on their commonalities (Tomlinson, 2001), it seems like a promising approach. This research project supports this idea, as, despite the heterogeneity of effects, a mainly positive impact of DI on high-achieving students' achievement and motivational-affective characteristics was found. This is in accordance with previous studies that found positive impacts on students in general (Deunk et al., 2018; Smale-Jacobse et al., 2019) as well as a current systematic review that summarized effective educational practices for high-achieving students in mixed-ability classrooms (Barbier et al., 2022).

Besides the empirical support for a mainly positive impact of DI on high-achieving students' outcomes, a further central added value is that different conditions for a successful implementation of DI for high-achieving students in mixed-ability classrooms were considered. It was found that, despite the positive appraisal of DI as a useful support measure for high-achieving students, this approach is not used proactively and regularly. Hence, the most often mentioned facilitator regarding DI for high-achieving students in the current research project, namely professional staff development, is discussed in more detail. It might help teachers to implement DI more often and therefore help to unfold its potential for high-achieving students in mixed-ability classrooms.

On the one hand, we know that gifted education is difficult and teachers cannot do it on their own without any support (Kahmann et al., 2022; Sisk, 2009). On the other hand, a survey in the United States showed that out of 48 participating states, only three states indicated that university courses in gifted education are mandatory for all preservice teachers (Rinn et al., 2020). In order to get a better picture how effective professional staff development in the context of DI is for teacher and student outcomes, a current meta-analysis summarized 27 studies (Kahmann et al., 2022). The main results showed that professional staff development significantly improved teachers' knowledge, attitudes and DI practices. However, no significant direct effect on students' learning outcomes was found. Thus, further research is needed how professional staff development should be designed to also affect students' outcomes (Kahmann et al., 2022). Another potential of professional staff development is to overcome teachers' misconceptions (e.g., high-achieving students do it on their own) through refutation texts that take up certain misconceptions and counter them through information from, for example, empirical studies (Lassonde et al., 2016; Menz et al., 2021). Thereby, not only high-achieving students in general, but also the different high-achieving student subgroups might be a good starting point in order to differentiate instruction appropriately for this heterogeneous group of students. So far, special interventions for the subgroup of Underachievers exist and were also investigated regarding their effectiveness (Steenbergen-Hu et al., 2020). For other high-achieving student subgroups, there is still a lack of research.

Below, the two main research fields of the dissertation framework are discussed from a broader educational perspective. Regarding the research field of classification, the question of whether it is even necessary to identify and classify high-achieving students or whether we should comply with James Borland (2005, p. 1), who claims for "Gifted Education Without Gifted Children" is discussed. Regarding the research field of individualized in-school support, from the perspective of educational justice, the question of whether we should claim for special education for high-achieving students, although it may widen the achievement gap, is addressed.

6.2 Classification of High-Achieving Students – Is it Really Necessary?

When reading through the giftedness literature, many different opinions occur on how high-achieving students should be identified and even if it is necessary to have a (common) definition of these individuals (Borland, 2005; Carman, 2013; Coleman, 2004; Merry, 2008). On the one hand, Borland (2005, p. 1), for example, stated that the aim is to have "gifted education without gifted children". He thereby claims that he advocates appropriate gifted education but that this does not need the identification and labeling of high-achieving students. Rather the consideration of all individual needs of students independent of their classification is needed. Through DI, the individual needs of all students, independent of their label as high-achieving or not-as-high-achieving identified, could be fostered (Borland, 2005). On the other hand, researchers claim that for the research field and the educational practice, the use of some definition of high-achieving students is necessary and even non-debatable (Coleman, 2004; Merry, 2008). Without a common definition, the coherence within this research field is questionable and might lead to non-generalizable results between different studies (Carman, 2013; Neuendorf et al., 2022).

A very broad but well-known and widely accepted definition is the one from the Marland report in the 1970s, stating that

> gifted and talented children are those identified by professionally qualified persons who by virtue of outstanding abilities, are capable of high performance. These are children who require differentiated educational programs and/or services beyond those normally provided by the regular school program in order to realize their contribution to self and society. (Marland, 1971, p. 8)

In an ideal world, all students are considered individually and are nurtured according to their individual needs at this exact moment. From a practical teacher perspective, this individualized consideration is very cumbersome and the reduction of heterogeneity through the classification of students is necessary in order to remain capable of acting and fostering the learning of the whole diverse student group (Howard & Hoffman, 2018; Trautmann & Wischer, 2011). In order to find a balance between the criticized simplistic dichotomization of high-achieving versus not-as-high-achieving identified students and the fully individualized perspective, the current research project detected different high-achieving student subgroups.

Thereby, clearly distinguishable subgroups of high-achieving students through both methodological approaches were found that differed regarding their achievement and motivational-affective characteristics. Due to their partially very different characteristics, they also require different support measures. Hence, the findings supported the statement that a dichotomization between high-achieving versus not-as-high-achieving identified students is too simplistic. Further, the methodological approach and the identification indicators used are sensitive to different characteristics of high-achieving students and can therefore influence which students are labeled as high-achieving in which domain. To sum up, knowledge that there are different high-achieving student subgroups and how they can differ from each other is gained. However, allocating students to these different subgroups is not easy and might differ depending on the method used. Researchers and practitioners should be sensitized regarding this knowledge.

6.3 Gifted Education—Praise or Malice for the Education System

With the knowledge of high-achieving student subgroups' strengths and weaknesses regarding their achievement and motivational-affective characteristics, a next step is to think about the appropriate educational promotion of these subgroups to serve their individual needs (Borland, 2005). With regard to the appropriate education for high-achieving students, one often heard criticism is that specific support measures for high-achieving students are elitist and unfair (Kunze et al., 2019; Merry, 2008; Plucker & Callahan, 2020). Fairness was often used as a synonym for equality, whereby the idea that fairness implies teaching all children in the same way (one-size-fits-all approach) occurred (Cooper, 2009). In contrast, fairness can also be seen as appropriate challenge for all students, independent of their ability. Thereby, appropriate challenge can be defined as the state where students receive tasks that enable individual growth regarding their abilities and get enough support in order to master these tasks (Merry, 2008). According to this definition and the opinion of different researchers, educational justice does *not* mean the same education for all students but the same chance for all students to get appropriate education (Borland, 2005; Moltzen, 2011) even though this might imply a widening achievement gap (Gagné, 2005). Thus, due to the neglect of high-achieving students with regard to promotion approaches in educational settings, first the added value and necessity of gifted education at an individual and a societal level are discussed. Second, the importance of considering high-achieving students' differences and not seeing them as one homogeneous group is underlined.

In contrast to the myth that gifted education is elitist and high-achieving students do it on their own, the Marland report even included in its definition of high-achieving students that they need DI beyond the regular instruction in school. High-achieving students also face challenges in the educational context and cannot withstand educational disregard (Marland, 1971; Moon, 2009). To reinforce this point, Borland (2005) argued that without DI, high-achieving students are among those who are let down the most. Appropriate education is as important for high-achieving students as it is for their classmates, as they also need support to develop their full (academic) potential. If high-achieving students are underchallenged for a long period of time, various negative consequences such as a decrease in their motivation and achievement might occur (KMK, 2015). Inappropriate education leads to high-achieving students' disengagement and that in turn causes them to stop performing. Thus, one recommendation is to provide DI with, for example, choice and challenge (Kanevsky & Keighley, 2003). If instruction contains meaningfulness and appropriate challenge, it is worth being motivated and pursuing (Little, 2012). "Hence, educational attainment or success in school is directly tied to the level of motivation children experience vis-a-vis the sorts of educational tasks and projects they are given" (Merry, 2008, p. 59).

Further, the promotion of high-achieving students could also benefit the people surrounding them and a society as a whole (Kahveci & Akgül, 2014; Merry, 2008; Ninkov, 2020). A higher percentage of high-achieving students could, for example, increase the educational achievement of a country (OECD, 2019b) and reduce skill shortages (Cappelli, 2015; European Centre for the Development of Vocational Training, 2016). Promotion of high-achieving students benefits the whole society, as the development of their potential might enable them to take leading roles in society (Plucker & Callahan, 2020). However, when promoting high-achieving students, one should consider that resources are limited, and that the promotion of high-achieving students should not be at the expense of their classmates.

Further, high-achieving students should not be considered as one homogeneous group (Ziernwald, Schiepe-Tiska, & Reiss, 2022) that should receive the same educational approach. Ideally, one should find instructional approaches that promote all learners independent of their ability level and individual characteristics (Merry, 2008). According to this consideration and the point that educational justice and fairness in the current dissertation equal the understanding of an appropriate challenge for all students according to their individual needs and *not* an equal one-size-fits-all approach, DI seems to be an appropriate educational strategy to promote high-achieving students (Barbier et al., 2022; Ziernwald, Hillmayr, & Holzberger, 2022), while taking into account their different strengths and weaknesses. DI could promote high-achieving students in mixed-ability classrooms without harming their classmates with differing ability levels, as this instructional approach tries to consider the individual needs of all students (Merry, 2008; Tomlinson, 2001). Accordingly, DI has the potential to foster all students in mixed-ability classrooms and at the same time also account for the heterogeneity within high-achieving students (Borland, 2005; Deunk et al., 2018; Puzio et al., 2020; Smale-Jacobse et al., 2019; Ziernwald, Hillmayr, & Holzberger, 2022).

In order to achieve the aim of an appropriate promotion of high-achieving students, a rethinking amongst researchers, practitioners, and policy-makers is necessary. This rethinking should consider, on the one hand, that high-achieving students are a student population that merit promotion and, on the other hand, that high-achieving students are not a homogeneous subgroup but have different strengths and weaknesses. To date, no clear legal regulations guiding gifted education exist. Therefore, we still need to argue for acknowledging the individual differences of high-achieving students and the necessity of promoting them according to their individual needs and potential. Appropriate promotion through DI is important to reach educational justice for the individual and to fully use the human resources of a society (KMK, 2015, 2016a; Ninkov, 2020).

6.4 Methodological Considerations

In this chapter, the used database and methods as well as the analytical decisions made in the current research project are discussed regarding their strengths and limitations. From this, recommendations for future research are derived.

6.4.1 Methodological Strengths

The current research project used a sophisticated database as well as methodological approaches, which helped to overcome methodological problems of previous research. In the following, the use of data from large-scale assessments such as PISA and NAS, and the creation of a systematic mixed-methods review are discussed.

One commen criticism in giftedness research is that the investigated samples are small in size (e.g., García-Martínez et al., 2021; Rost, 1991; Ziernwald, Hillmayr, & Holzberger, 2022) and often very selective as study participants are part of special gifted programs. One possibility to evade these criticisms is to select high-achieving students from a large unselected sample in regular mixed-ability classrooms (Preckel et al., 2008). This was done in the current research project as data from two large-scale assessments were used to answer the research questions about high-achieving students' classification and characterization. The sample of high-achieving students was selected from a representative sample of 9th graders in Germany. As a result, the sample (N = 1563) was large for research in gifted education, including high-achieving students found in regular mixed-ability classrooms. With this sample, the criticism to have a small and highly selective group of high-achieving students could be overcome. Thus, the results can be generalized and applied to a broader student population.

Further, various achievement indicators that might not be comparable are used across studies to identify and classify high-achieving students (Carman, 2013; Neuendorf et al., 2022). School grades, for example, which were used for the identification of high-achieving students in many previous studies (Neuendorf et al., 2022), are dependent on the context and, therefore, hardly comparable across different classes and schools (Bayer & Zinn, 2018) and therefore studies. In contrast, the current research project used standardized achievement indicators, which can be compared across different grades, school types, and partially even across different countries (OECD, 2013a; Pant et al., 2013).

Another strength of the current research project is the performance of a systematic mixed-methods review, as it unites the potential of systematic research syntheses and mixed-methods approaches. The most obvious advantage is that research syntheses help to get an overview of the existing research and at the same time disclose research gaps in a certain research field, such as giftedness research. In recent years, this aspect has gained importance as the number of publications has increased (Beelmann, 2014; Holzberger & Ziernwald, 2020). Especially for practitioners, this advantage is highly relevant, as a perceived lack of time and sourcing skills positively influenced teachers' irrelevance perceptions of research to their educational practice (Thomm et al., 2021). Thus, research syntheses can help decrease the researcher-practitioner-gap (Beelmann, 2014; Holzberger & Ziernwald, 2020). Further, through taking into account various primary studies, the results are based on a larger overall sample size. This advantage is especially helpful in giftedness research as it is often based on small samples (Rost, 1991). In addition to the general potential of research syntheses, mixed-methods research syntheses enable further advantages for giftedness research. By considering systematically qualitative, quantitative, and mixed-methods primary studies, research questions could be answered more comprehensively, specifically, and detailed. On the one hand, the strength of quantitative designs and, on the other hand, the strength of qualitative designs can be used. Additionally, the different research approaches can compensate for their respective weaknesses (Heyvaert et al., 2013). With these advantages, the limitations of previous research syntheses on DI, stating that both quantitative and qualitative studies should be taken into account to gain an even more comprehensive picture of DI (Smale-Jacobse et al., 2019), could be overcome.

6.4.2 Limitations and Directions for Future Research

Despite the methodological strengths and the findings, which close research gaps and have crucial practical implications, there are also limitations, which should be discussed and lead to directions for future research. First, Renzulli's Three-Ring Conception of Giftedness (Renzulli, 2005) mentions creativity, next to achievement and motivational-affective characteristics, as worth investigating when characterizing high-achieving students. Accordingly, it was repeatedly claimed that the interplay between creativity, achievement, and intelligence should be discussed. Creativity can help contribute to innovations for society and is linked to many student characteristics such as intrinsic motivation and academic achievement (Gutman & Schoon, 2013; Hennessey, 2017; Renzulli, 2017). The current research project focused on different achievement indicators for the classification of high-achieving students into subgroups, as these are the most often used operationalizations of high-achieving students (Carman, 2013; Neuendorf et al., 2022; Ziernwald, Hillmayr, & Holzberger, 2022). However, with creative thinking being the innovative domain in PISA 2022 (OECD, 2019c), data from PISA

2022 with mathematics again as the main domain will offer a great opportunity to combine the strength of standardized achievement indicators for mathematical competence and figural reasoning of the current research project and the objective assessment of creative thinking for identifying and characterizing high-achieving student subgroups. Further, this data would also make it possible to expand the current findings by including, in addition to mathematics as a domain-specific indicator, other domains such as reading. Research on ability tilt, defined as difference between mathematics and verbal achievement within an individual, showed that this tilt is associated with individuals' educational and career choices (Coyle et al., 2014; Sadowski & Zawistowska, 2020). Hence, the inclusion of different domain-specific indicators while investigating high-achieving students' strengths might be useful.

Second, cross-sectional data to identify and characterize high-achieving student subgroups were used. Based on previous research regarding the stability of the investigated achievement indicators (e.g., Holenstein et al., 2021; Larsen et al., 2008; Yu et al., 2018), one can reasonably assume that the subgroups detected are quite stable throughout the school career. Despite this evidence for stability, there is research that points towards another interesting development of high-achieving students' achievement. Their achievement is often compared to other not-as-high-achieving identified students. Thereby, research found inconsistent results if a Matthew effect, a compensatory effect, or a stable pattern exists (e.g., Baumert et al., 2012; Pfost et al., 2014). A recent study by Neuendorf et al. (2020) found support for the compensatory effect for high-achieving students in mathematics and reading from grades 5 to 9. Hence, one implication of the authors was that future studies should focus on applying DI for high-achieving students and investigate if this influences high-achieving students' achievement trajectories differently (Neuendorf et al., 2020). In order to have a closer look at the stability of high-achieving student subgroups, further research might benefit from latent transition analysis. Here, class probabilities for the initial measurement and probabilities for the transition from the initial classes to classes to a later measurement time are defined (Hickendorff et al., 2018). With this model, one could investigate if students move between the identified subgroups of high-achieving students and, if yes, the transitions between which subgroups are most common. This additional information might be interesting itself and also for the promotion of high-achieving students. Do children show similar individual needs once identified or do the characteristics and individual needs change over time? If teachers are aware of different subgroups and student transitions between them, it can facilitate appropriate DI in their lessons (Schlatter et al., 2021).

The third limitation of the current research project is the fact that the impact of DI was investigated on high-achieving students in general (Ziernwald, Hillmayr, & Holzberger, 2022) and not on specific high-achieving student subgroups. This can be seen as problematic as one central conclusion of this dissertation is that high-achieving students are not a homogeneous group and their individual differences should be considered for gifted education (Ziernwald, Schiepe-Tiska, & Reiss, 2022). The summarized primary studies in the current research project used different identification approaches (i.e., standardized tests, teacher recommendations, or multiple methods) that may have resulted in different high-achieving student subgroups. However, the number of included studies was not large enough to make specific moderator analyses based on the different detected subgroups. Hence, the specific impact of educational interventions on high-achieving student subgroups is still an important research gap that should be investigated in the future. A meta-analysis, for example, found that the differences in the perceived academic competence and the global self-concept between high-achieving and not-as-high-achieving identified students varied, depending on the identification method used for high-achieving students (Litster & Roberts, 2011). This result is in accordance with the current research project that claims that the identification method used results in different subgroups of high-achieving students, which may differ from each other extensively.

Fourth, in accordance with previous reviews in the field of giftedness education (e.g., Barbier et al., 2022; García-Martínez et al., 2021) and the third limitation, a lack of studies with high methodological standards was found and studies widely differed regarding their conceptualizations of the DI approaches and the identification of high-achieving students. Accordingly, more primary studies with high methodological standards in the field of giftedness research with practical implications for gifted education are needed. With this additional knowledge, future reviews may be able to explain which specific aspects of DI are particularly effective for which high-achieving student subgroup and further elucidate the heterogeneous impacts found. Hence, this research project detected key research gaps that provide important starting points for future research in order to advance the field of giftedness research in a targeted manner.

6.5 Practical Implications

In addition to implications for giftedness research, practical implications for the education of high-achieving students are provided. Both studies in the current research project focused on a broad sample of high-achieving students, wherefore the results are generalizable to a broader population and not just to a very specific small group (e.g., students with an IQ score higher than 130). The heterogeneity amongst this broader high-achieving student sample and that DI could be an appropriate approach towards fostering them according to their individual needs was demonstrated. Therefore, one main aim of this dissertation is to sensitize researchers, politicians, and especially teachers regarding the heterogeneity of high-achieving students. The achievement and motivational-affective characteristics play an important role in taking their individual needs for education into account in order to help them fulfill their full potential.

This sensitization is one important prerequisite for the development of teachers' professional competences for the education of high-achieving students. Teachers' beliefs, values, and motivational orientations about DI are crucial success factors for its implementation (Jäger, 2019). The results of the current research project show that teachers and students perceived DI for high-achieving students as useful and that their beliefs were quite positive. On the other hand, sensitization is not enough to systematically develop professional competences (Kunze et al., 2019). When German teachers were asked if approaches for individualized learning were part of their teacher training, more than a third of the responding teachers said no, and less than two-thirds said yes (Mang et al., 2021). Therefore, following one of the findings of the systematic mixed-methods review, professional development for student teachers and teachers is an essential field of action.

However, it is also crucial to remember that individualized promotion of high-achieving students and students in general cannot be the task of one teacher alone. On the one hand, teachers are essential stakeholders, and without them, the implementation of individualized learning environments is impossible. However, on the other hand, this implementation could only succeed if the school as a whole supports this task and the individual teacher (Arndt et al., 2014; Jäger, 2019). This is in accordance with the aspect of leadership support that was found as one facilitating factor for implementing DI for high-achieving students in mixed-ability classrooms. Only if the importance of the promotion is paired with the availability of the corresponding material resources and teacher competences in this regard, the promotion of high-achieving students through DI can succeed.

One further conceivable way of supporting the classroom teacher in inclusive educational settings is to work in multiprofessional teams as is known from inclusive education with disadvantaged children (Wocken, 2019). One five-layered model reflecting the inclusive education in schools indicates students with their individual needs as the innermost layer. Therefore, the heterogeneity of students and the knowledge about their respective strengths and weaknesses is essential. This innermost layer is followed by inclusive instruction such as through DI taking students' differences into account. The third layer contains multiprofessional teams (Gebhardt & Heimlich, 2018), which are seen as key characteristic of inclusive education (Wocken, 2019) and good inclusive schools (Arndt et al., 2014). Different cooperations through multiprofessional teams are conceivable. On the one hand, team teaching with two teachers who conduct the instruction for all students is possible. On the other hand, special needs teachers, school psychologists or learning therapists can support the leading teacher through supporting individual students in the classroom (Verband Bildung und Erziehung, 2016). One further advantage of teaching in multiprofessional teams, in addition to the provision of relief, is the opportunity to reflect on one's own instruction, which is of high importance, especially for inclusive education (Arndt et al., 2014). Thus, multiprofessional teams should not only include experts for disadvantaged students, but also experts in gifted education in order to also meet high-achieving students' needs and ideally foster their potential. The fourth layer of the inclusive education model relates to school conceptions and school life, and the fifth layer to networking with the school environment (Gebhardt & Heimlich, 2018).

Another option that can help relieve teachers and still support high-achieving students appropriately according to their potential is the use of digital media. One crucial advantage of digital media is the opportunity to individualize instruction (Leutner, 2002). Research syntheses summarizing the effects of intelligent tutoring systems, which are characterized by the fact that they adapt to the students' individual requirements and provide individualized instruction, showed positive effects on students' achievement compared to instructor-led large-group instruction, individual work with textbooks, as well as learning without digital media. Positive effects were found across different grades and subjects (Hillmayr et al., 2020; Ma et al., 2014). Further, computer-based DI positively affected students cognitive achievement (Deunk et al., 2018). However, there is still a lack of research regarding the effectiveness of computer-based DI for high-achieving students and the use of digital media for subgroups of high-achieving students (Deunk et al., 2018; Periathiruvadi & Rinn, 2012). Nevertheless, this research project provided further hints that computer-based DI is also useful for high-achieving students (Ziernwald, Hillmayr, & Holzberger, 2022).

Last but not least, extracurricular measures—although they were not in the focus of the current dissertation—should be seen as an important complement to the promotion of high-achieving students in inclusive educational settings (e.g., García-Martínez et al., 2021; Olszewski-Kubilius, 2009). Thereby, mentoring can be seen as one very successful measure in gifted education that could especially influence the interest in a certain field, such as STEM (Stoeger et al., 2016; Stoeger et al., 2017). This positive impact on motivational-affective characteristics can be considered at least as relevant as a positive impact on achievement. For example, in the context of skills shortages, motivational-affective characteristics in particular can influence which field of study one chooses (e.g., Taskinen et al., 2013). Hence, the detected high-achieving students' heterogeneity is not only relevant for inclusive educational settings but also for extracurricular measures. Depending on the high-achieving students' individual potential and needs, the support measure in the educational context, including support measures within educational settings as well as extracurricular approaches, should be considered.

7 Conclusion

The classification and individualized in-school support of high-achieving students in educational contexts has gained attention in recent years. This dissertation addressed these two main research fields and thereby tackled repeatedly mentioned challenges from research and practice. In summary, the dissertation helps to counter long-held misconceptions from laypeople and academics regarding high-achieving students (e.g., high-achieving students are gifted in all areas, high-achieving students do it on their own) in educational contexts. A multidimensional perspective including both domain-general and domain-specific achievement indicators as well as motivational-affective characteristics is necessary for identifying the heterogeneous characteristics and needs of high-achieving students in an appropriate manner. By addressing high-achieving students' needs and by providing individualized in-school support through DI in inclusive educational settings, high-achieving students receive the attention they deserve as a long-neglected student group. With this individualized and inclusive approach, the labeling of high-achieving students can be de-emphasized and the promotion programs can be designed more inclusive for a broader student population. Hence, this dissertation helps to realize the full potential of high-achieving and potentially high-achieving students and thus helps to promote one of the most important human resources of a society.

References

Acar, S., Sen, S., & Cayirdag, N. (2016). Consistency of the performance and nonperformance methods in gifted identification. *Gifted Child Quarterly*, *60*(2), 81–101. https://doi.org/10.1177/0016986216634438

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

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, *61*(3), 315–343. https://doi.org/10.3102/00346543061003315

Amor, A. M., Hagiwara, M., Shogren, K. A., Thompson, J. R., Verdugo, M. Á., Burke, K. M., & Aguayo, V. (2019). International perspectives and trends in research on inclusive education: A systematic review. *International Journal of Inclusive Education*, *23*(12), 1277–1295. https://doi.org/10.1080/13603116.2018.1445304

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

Armstrong, F. (2011). Inclusive education: School cultures, teaching and learning. In G. Richards & F. Armstrong (Eds.), *Teaching and learning in diverse and inclusive classrooms: Key issues for new teachers* (pp. 7–18). Routledge.

Armstrong, F., & Richards, G. (2011). Introduction. In G. Richards & F. Armstrong (Eds.), *Teaching and learning in diverse and inclusive classrooms: Key issues for new teachers* (pp. 1–6). Routledge.

Arndt, A.-K., Stenger, S., & Werning, R. (2014). Gestaltung und Entwicklung inklusiven Unterrichts [Designing and developing inclusive instruction]. In A.-K. Arndt, A. Harting, P. Katzer, M. Laubner, S. Stenger, & R. Werning (Eds.), *Schulmanagement-Handbuch: Vol. 152. Inklusiver Unterricht: Leitideen zur Organisation und Kooperation* (pp. 6–24). Oldenbourg.

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

Baumert, J., Brunner, M., Lüdtke, O., & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? - Resultate kumulativer Wissenserwerbsprozesse [What do international school achievement studies measure? - Results of cumulative knowledge acquisition processes]. *Psychologische Rundschau*, *58*(2), 118–128. https://doi.org/10.1026/0033-3042.58.2.118

Baumert, J., Nagy, G., & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, *83*(4), 1347–1367. https://doi.org/10.1111/j.1467-8624.2012.01779.x

Bayer, M., & Zinn, S. (2018). *Zur Entwicklung des Zusammenhangs und der Erklärbarkeit von Leistungen und Kompetenzen bei Schülerinnen und Schülern der Sekundarstufe I*. LIfBi Working Paper No. 77 [On the development of the correlation and the explanability of grades and competencies among pupils in secondary schools – An explorative study]. Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Beelmann, A. (2014). Möglichkeiten und Grenzen systematischer Evidenzkumulation durch Forschungssynthesen in der Bildungsforschung [Potentials and limits of the systematic accumulation of evidence via systematic research synthesis within educational research]. *Zeitschrift für Erziehungswissenschaft*, *17*(Supplemental), 55–78. https://doi.org/10.1007/s11618-014-0509-2

Bogat, G. A., von Eye, A., & Bergman, L. R. (2016). Person-oriented approaches. In D. Cicchetti (Ed.), *Developmental psychopathology* (pp. 797–845). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119125556.devpsy118

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.

Bond, L. A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research, and Evaluation*, *5*, Article 2. https://doi.org/10.7275/dy7r-2x18

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

Bönsch, M. (2004). *Differenzierung in Schule und Unterricht: Ansprüche, Formen, Strategien* [Differentiation in school and teaching: Demands, forms, strategies]. Oldenbourg.

Booth, T., & Ainscow, M. (2002). *Index for inclusion: Developing learning and partcipation in schools.* Centre for Studies on Inclusive Education. https://www.eenet.org.uk/resources/docs/Index%20English.pdf

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. https://doi.org/10.1017/CBO9780511610455.002

Cappelli, P. H. (2015). Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the United States. *ILR Review*, *68*(2), 251–290. https://doi.org/10.1177/0019793914564961

Carman, C. A. (2013). Comparing apples and oranges: Fifteen years of definitions of giftedness in research. *Journal of Advanced Academics*, *24*(1), 52–70. https://doi.org/10.1177/1932202X12472602

Castejón, J. L., Gilar, R., Miñano, P., & González, M. (2016). Latent class cluster analysis in exploring different profiles of gifted and talented students. *Learning and Individual Differences*, *50*, 166–174. https://doi.org/10.1016/j.lindif.2016.08.003

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*(3), 153–193. https://doi.org/10.1037/h0059973

Civitillo, S., Denessen, E., & Molenaar, I. (2016). How to see the classroom through the eyes of a teacher: Consistency between perceptions on diversity and differentiation practices. *Journal of Research in Special Educational Needs*, *16*(s1), 587–591. https://doi.org/10.1111/1471-3802.12190

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.

Coleman, L. J. (2004). Is consensus on a definition in the field possible, desirable, necessary? *Roeper Review*, *27*(1), 10–11. https://doi.org/10.1080/02783190409554280

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Wiley.

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.

Cooper, C. R. (2009). Myth 18: It is fair to teach all children the same way. *Gifted Child Quarterly*, *53*(4), 283–285. https://doi.org/10.1177/0016986209346947

Coyle, T. R., Purcell, J. M., Snyder, A. C., & Richmond, M. C. (2014). Ability tilt on the SAT and ACT predicts specific abilities and college majors. *Intelligence*, *46*, 18–24. https://doi.org/10.1016/j.intell.2014.04.008

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. (2017). Maximizing potential: A school-based conception of psychosocial development. *High Ability Studies*, *28*(1), 43–58. https://doi.org/10.1080/13598139.2017.1292896

Dai, D. Y., Swanson, J. A., & Cheng, H. (2011). State of research on giftedness and gifted education: A survey of empirical studies published during 1998-2010 (April). *Gifted Child Quarterly*, *55*(2), 126–138. https://doi.org/10.1177/0016986210397831

Davis, G. A. (2009). New developments in gifted education. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 1035–1044). Springer.

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

Deutsch, M. (1985). *Distributive justice: A social-psychological perspective*. Yale University Press.

Dixson, D. D., Peters, S. J., Makel, M. C., Jolly, J. L., Matthews, M. S., Miller, E. M., Rambo-Hernandez, K. E., Rinn, A. N., Robins, J. H., & Wilson, H. E. (2020). A call to reframe gifted education as maximizing learning. *Phi Delta Kappan*, *102*(4), 22–25. https://doi.org/10.1177/0031721720978057

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology*, *7*, Article 508. https://doi.org/10.3389/fpsyg.2016.00508

Dreeszen, J. L. (2009). *The impact of differentiation on the critical thinking of gifted readers and the evolving perspective of the fifth grade classroom teacher* [Doctoral dissertation]. Kansas State University, Manhatten. 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*, 249–277. https://doi.org/10.1007/s11618-018-0840-0

Ehmke, T., Köller, O., & Stanat, P. (2017). Äquivalenz der Erfassung mathematischer Kompetenzen in PISA 2012 und im IQB-Ländervergleich 2012 [Equivalence in measuring mathematical competence in PISA 2012 and in the National Assessment 2012]. *Zeitschrift für Erziehungswissenschaft*, *20*(S2), 37–59. https://doi.org/10.1007/s11618-017-0751-5

Ehmke, T., van den Ham, A.-K., Sälzer, C., Heine, J., & Prenzel, M. (2020). Measuring mathematics competence in international and national large scale assessments: Linking PISA and the National Educational Panel Study in Germany. *Studies in Educational Evaluation*, *65*, Article 100847. https://doi.org/10.1016/j.stueduc.2020.100847

European Centre for the Development of Vocational Training. (2016). *Skill shortage and surplus occupations in Europe: Cedefop insights into which occupations are in high demand - and why*. https://epale.ec.europa.eu/sites/default/files/skill_shortage_and_surplus_occupations_in_europe.pdf

Fischer, C. (2019). Professionalisierung von Lehrpersonen zur individuellen Begabungsförderung [Professionalization of teachers for the individualized promotion of giftedness]. In C. Reintjes, I. Kunze, & E. Ossowski (Eds.), *Begabungsförderung und Professionalisierung: Befunde, Perspektiven, Herausforderungen* (pp. 174–189). Julius Klinkhardt.

Fischer, C., & Müller, K. (2014). Gifted education and talent support in Germany. *CEPS Journal*, *4*(3), 31–54. https://doi.org/10.25656/01:9616

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies*, *15*(2), 119–147. https://doi.org/10.1080/1359813042000314682

Gagné, F. (2005). From noncompetence to exceptional talent: Exploring the range of academic achievement within and between grade levels. *Gifted Child Quarterly*, *49*(2), 139–153. https://doi.org/10.1177/001698620504900204

García-Martínez, I., Gutiérrez Cáceres, R., La Luque de Rosa, A., & León, S. P. (2021). Analysing educational interventions with gifted students. Systematic review. *Children*, *8*(5), Article 365. https://doi.org/10.3390/children8050365

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

Gebhardt, M., & Heimlich, U. (2018). Inklusion und Bildung [Inclusion and education]. In R. Tippelt & B. Schmidt-Hertha (Eds.), *Handbuch Bildungsforschung* (pp. 1241–1260). Springer.

General Assembly of the United Nations. (2019). Transforming our world: The 2030 Agenda for Sustainable Development. In J. I. Walker, A. Pekmezovic, & G. Walker (Eds.), *Sustainable Development Goals: Harnessing business to achieve the SDGs through finance, technology and law reform* (pp. 333–374). Wiley.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, *18*(8), 519–521. https://doi.org/10.1037/h0049294

Grimm, K. J., Mazza, G. L., & Davoudzadeh, P. (2017). Model selection in finite mixture models: A k-fold cross-validation approach. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(2), 246–256. https://doi.org/10.1080/10705511.2016.1250638

Gubbels, J., Segers, E., & Verhoeven, L. (2018). How children's intellectual profiles relate to their cognitive, socio-emotional, and academic functioning. *High Ability Studies*, *29*(2), 149–168. https://doi.org/10.1080/13598139.2018.1507902

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). *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

Heller, K. A. (2001). Teil I: Projektziele, Untersuchungsergebnisse und praktische Konsequenzen [Part I: Project objectives, study results, and practical consequences]. In K. A. Heller (Ed.), *Hochbegabung im Kindes- und Jugendalter* (2nd ed., pp. 21–40). Hogrefe.

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. https://doi.org/10.1017/CBO9780511610455.010

Hennessey, B. A. (2017). Intrinsic motivation and creativity in the classroom: Have we come full circle? In R. A. Beghetto & J. C. Kaufman (Eds.), *Nurturing creativity in the classroom* (2nd ed., pp. 227–264). Cambridge University Press.

Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: Definition, framework, and potential. *Quality & Quantity*, *47*(2), 659–676. https://doi.org/10.1007/s11135-011-9538-6

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences*, *66*, 4–15. https://doi.org/10.1016/j.lindif.2017.11.001

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, *153*, Article 103897. https://doi.org/10.1016/j.compedu.2020.103897

Hofer, S. I., & Stern, E. (2016). Underachievement in physics: When intelligent girls fail. *Learning and Individual Differences*, *51*, 119–131. https://doi.org/10.1016/j.lindif.2016.08.006

Holenstein, M., Bruckmaier, G., & Grob, A. (2021). Transfer effects of mathematical literacy: An integrative longitudinal study. *European Journal of Psychology of Education*, *36*, 799–825. https://doi.org/10.1007/s10212-020-00491-4

Holzberger, D., & Ziernwald, L. (2020). Forschungsflut - den Überblick behalten! [Research flood - keeping the overview!]. *Pädagogische Führung*, *1*, 25–27.

Hong, E., & Aqui, Y. (2004). Cognitive and motivational characteristics of adolescents gifted in mathematics: Comparisons among students with different types of giftedness. *Gifted Child Quarterly*, *48*(3), 191–201. https://doi.org/10.1177/001698620404800304

Howard, M. C., & Hoffman, M. E. (2018). Variable-centered, person-centered, and person-specific approaches. *Organizational Research Methods*, *21*(4), 846–876. https://doi.org/10.1177/1094428117744021

Jäger, S. (2019). Individuelle Förderung - Anspruch und Realität in der schulpraktischen Umsetzung [Individualized Promotion - Demand and reality in the practical implementation in schools]. In C. Drüke-Noe, U. Hauck-Thum, H. Hoppe, S. Jäger, U. Maier, & K. Metz (Eds.), *Schulmanagement-Handbuch: Vol. 174. Individuelle Förderung in der Schule: Grundlagen, Konzepte, Praxis* (pp. 6–23). Cornelsen.

Kahmann, R., Droop, M., & Lazonder, A. W. (2022). Meta-analysis of professional development programs in differentiated instruction. *International Journal of Educational Research*, *116*, Article 102072. https://doi.org/10.1016/j.ijer.2022.102072

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., & Keighley, T. (2003). To produce or not to produce? Understanding boredom and the honor in underachievement. *Roeper Review*, *26*(1), 20–28. https://doi.org/10.1080/02783190309554235

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, 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

Krämer, S., Möller, J., & Zimmermann, F. (2021). Inclusive education of students with general learning difficulties: A meta-analysis. *Review of Educational Research*, *91*(3), 432–478. https://doi.org/10.3102/0034654321998072

Kulik, J. A., & Kulik, C.-L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly*, *36*(2), 73–77. https://doi.org/10.1177/001698629203600204

Kultusministerkonferenz. (2009). *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. (2010). *Förderstrategie für leistungsschwächere Schülerinnen und Schüler* [Support strategy for lower-achieving students]. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_03_04-Foerderstrategie-Leistungsschwaechere.pdf

Kultusministerkonferenz. (2015). *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

Kultusministerkonferenz. (2016a). *Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler* [Joint initiative of the Federal Ministry of Education and Research and the Standing Conference of Ministers of Education and Cultural Affairs to promote high-achieving and potentially high-achieving students]. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2016/2016-11-28_Gem.Initiative_Leistungsstarke-Beschluss.pdf

Kultusministerkonferenz. (2016b). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring* [Overall strategy of the Standing Conference of Ministers of Education and Cultural Affairs on educational monitoring]. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf

Kunze, I., Reintjes, C., & Ossowski, E. (2019). Begabungsförderung und Professionalisierung im Spiegel des Diskurses um Bildungsgerechtigkeit [Promoting giftedness and professionalization in the mirror of the discourse on educational justice]. In C. Reintjes, I. Kunze, & E. Ossowski (Eds.), *Begabungsförderung und Professionalisierung: Befunde, Perspektiven, Herausforderungen* (pp. 217–224). Julius Klinkhardt.

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle-age. *Intelligence*, *36*(1), 29–34. https://doi.org/10.1016/j.intell.2007.01.001

Lassonde, K. A., Kendeou, P., & O'Brien, E. J. (2016). Refutation texts: Overcoming psychology misconceptions that are resistant to change. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 62–74. https://doi.org/10.1037/stl0000054

Lazarides, R., Dietrich, J., & Taskinen, P. H. (2019). Stability and change in students' motivational profiles in mathematics classrooms: The role of perceived teaching. *Teaching and Teacher Education*, *79*, 164–175. https://doi.org/10.1016/j.tate.2018.12.016

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin Company.

Leikin, R., Leikin, M., Paz-Baruch, N., Waisman, I., & Lev, M. (2017). On the four types of characteristics of super mathematically gifted students. *High Ability Studies*, *28*(1), 107–125. https://doi.org/10.1080/13598139.2017.1305330

Leutner, D. (2002). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme [Adaptivity and adaptability of multimedia teaching and information systems]. In L. J. Issing & P. Klimsa (Eds.), *Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxis* (3rd ed., pp. 115–125). Beltz PVU.

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

Litster, K., & Roberts, J. (2011). The self-concepts and perceived competencies of gifted and non-gifted students: A meta-analysis. *Journal of Research in Special Educational Needs*, *11*(2), 130–140. https://doi.org/10.1111/j.1471-3802.2010.01166.x

Little, C. A. (2012). Curriculum as motivation for gifted students. *Psychology in the Schools*, *49*(7), 695–705. https://doi.org/10.1002/pits.21621

Lohman, D. F., Gambrell, J., & Lakin, J. (2008). The commonality of extreme discrepancies in the ability profiles of academically gifted students. *Psychology Science Quarterly*, *50*(2), 269–282.

Lok, B., McNaught, C., & Young, K. (2016). Criterion-referenced and norm-referenced assessments: Compatibility and complementarity. *Assessment & Evaluation in Higher Education*, *41*(3), 450–465. https://doi.org/10.1080/02602938.2015.1022136

Loveless, T. (2008). Analysis of NAEP Data. In Thomas B. Fordham Institute (Ed.), *High-achieving students in the era of NCLB* (pp. 13–48). https://files.eric.ed.gov/fulltext/ED501703.pdf

Lucito, L. J. (1964). Gifted Children. In L. M. Dunn (Ed.), *Exceptional children in the schools* (pp. 179–238). Holt, Rinehart and Winston.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901–918. https://doi.org/10.1037/a0037123

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. https://doi.org/10.1037//1082-989X.7.1.19

Maker, C. J. (1996). Identification of gifted minority students: A national problem, needed changes and a promising solution. *Gifted Child Quarterly*, *40*(1), 41–50. https://doi.org/10.1177/001698629604000106

Maker, C. J., & Schiever, S. W. (2010). *Curriculum development and teaching strategies for gifted learners* (3rd ed.). PRO-ED.

Mammadov, S., Ward, T. J., Cross, J. R., & Cross, T. L. (2016). Use of latent profile analysis in studies of gifted students. *Roeper Review*, *38*(3), 175–184. https://doi.org/10.1080/02783193.2016.1183739

Mang, J., Seidl, L., Schiepe-Tiska, A., Tupac-Yupanqui, A., Ziernwald, L., Doroganova, A., Weis, M., Diedrich, J., Heine, J.-H., Gonzaléz Rodríguez, E., & Reiss, K. (2021). *PISA 2018 Skalenhandbuch: Dokumentation der Erhebungsinstrumente* [PISA 2018 Questionnaire handbook - Documentation of the assessed instruments]. Waxmann. https://doi.org/10.31244/9783830994961

Marland, S. P. (1971). *Education of the gifted and talented - Volume 1: Report to the Congress of the United States by the U. S. Commissioner of Education*. Office of Education. https://files.eric.ed.gov/fulltext/ED056243.pdf

McBee, M. T., Peters, S. J., & Waterman, C. (2014). Combining scores in multiple-criteria assessment systems. *Gifted Child Quarterly*, *58*(1), 69–89. https://doi.org/10.1177/0016986213513794

Menz, C., Spinath, B., & Seifried, E. (2021). Misconceptions die hard: Prevalence and reduction of wrong beliefs in topics from educational psychology among preservice teachers. *European Journal of Psychology of Education*, *36*, 477–494. https://doi.org/10.1007/s10212-020-00474-5

Merry, M. S. (2008). Educational justice and the gifted. *Theory and Research in Education*, *6*(1), 47–70. https://doi.org/10.1177/1477878507086730

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

Moltzen, R. (2011). Inclusive education and gifted and talented provision. In G. Richards & F. Armstrong (Eds.), *Teaching and learning in diverse and inclusive classrooms: Key issues for new teachers* (pp. 102–112). Routledge.

Moon, S. M. (2009). Myth 15: High-ability students don't face problems and challenges. *Gifted Child Quarterly*, *53*(4), 274–276. https://doi.org/10.1177/0016986209346943

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

Neidorf, T. S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing mathematics content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 assessments: Technical report* (NCES 2006-029). National Center for Education Statistics. https://nces.ed.gov/pubs2006/2006029.pdf

Neuendorf, C., Jansen, M., & Kuhl, P. (2020). Competence development of high achievers within the highest track in German secondary school: Evidence for Matthew effects or compensation? *Learning and Individual Differences*, *77*, Article 101816. https://doi.org/10.1016/j.lindif.2019.101816

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*, 280–289. https://doi.org/10.1007/s41134-020-00133-1

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). Yale University Press.

OECD. (2013a). *PISA 2012 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/9789264190511-en

OECD. (2013b). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs (Volume III)*. OECD Publishing. https://doi.org/10.1787/9789264201170-en

OECD. (2014). *PISA 2012 results: What students know and can do - Student performance in mathematics, reading and science (Volume I)* (Revised Edition). OECD Publishing. https://doi.org/10.1787/9789264208780-en

OECD. (2016). *PISA 2015 results: Excellence and equity in education (Volume I)*. OECD Publishing. https://doi.org/10.1787/9789264266490-en

OECD. (2018). *PISA 2022 mathematics framework (Draft)*. https://pisa2021-maths.oecd.org/files/PISA%202021%20Mathematics%20Framework%20Draft.pdf

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

OECD. (2019b). *PISA 2018 results: What students know and can do (Volume I)*. OECD Publishing. https://doi.org/10.1787/5f07c754-en

OECD. (2019c). *PISA 2021 Creative thinking framework (Third draft)*. https://www.oecd.org/pisa/publications/PISA-2021-creative-thinking-framework.pdf

OECD. (2020). *PISA 2024: Strategic vision and direction for science*. https://www.oecd.org/pisa/publications/PISA-2024-Science-Strategic-Vision-Proposal.pdf

Olszewski-Kubilius, P. (2009). Special schools and other options for gifted STEM students. *Roeper Review*, *32*(1), 61–70. https://doi.org/10.1080/02783190903386892

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (Eds.). (2013). *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* [The IQB National Assessment Study 2012: Competencies in mathematics and the sciences at the end of secondary level I]. Waxmann.

Periathiruvadi, S., & Rinn, A. N. (2012). Technology in gifted education: A review of best practices and empirical research. *Journal of Research on Technology in Education*, *45*(2), 153–169. https://doi.org/10.1080/15391523.2012.10782601

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research*, *84*(2), 203–244. https://doi.org/10.3102/0034654313509492

Plucker, J. A., & Callahan, C. M. (2020). The evidence base for advanced learning programs. *Phi Delta Kappan*, *102*(4), 14–21. https://doi.org/10.1177/0031721720978056

Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly*, *52*(2), 146–159. https://doi.org/10.1177/0016986208315834

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

Reinhold, F., Reiss, K., Diedrich, J., Hofer, S., & Heinze, A. (2019). Mathematische Kompetenz in PISA 2018 - aktueller Stand und Entwicklung [Mathematical literacy in PISA 2018 - current status and development]. In K. Reiss, M. Weis, E. Klieme, & O. Köller (Eds.), *PISA 2018: Grundbildung im internationalen Vergleich* (pp. 187–209). Waxmann. https://doi.org/10.31244/9783830991007

Reis, S. M., & Renzulli, J. S. (2009). Myth 1: The gifted and talented constitute one single homogeneous group and giftedness is a way of being that stays in the person over time and experiences. *Gifted Child Quarterly*, *53*(4), 233–235. https://doi.org/10.1177/0016986209346824

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

Renzulli, J. S. (2005). The Three-Ring Conception of Giftedness: A developmental model for promoting creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 246–279). Cambridge University Press. https://doi.org/10.1017/CBO9780511610455.015

Renzulli, J. S. (2017). Developing creativity across all areas of the curriculum. In R. A. Beghetto & J. C. Kaufman (Eds.), *Nurturing creativity in the classroom* (2nd ed., pp. 23–44). Cambridge University Press. https://doi.org/10.1017/9781316212899.006

Retelsdorf, J., Becker, M., Köller, O., & Möller, J. (2012). Reading development in a tracked school system: A longitudinal study over 3 years using propensity score matching. *British Journal of Educational Psychology*, *82*(4), 647–671. https://doi.org/10.1111/j.2044-8279.2011.02051.x

Rieser, S., Stahns, R., Walzeburg, A., & Wendt, H. (2016). Einblicke in die Gestaltung des Mathematik- und Sachunterrichts [Insights into the design of mathematics and general science instruction]. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert, D. Kasper, & Waxmann Verlag GmbH (Eds.), *TIMSS 2015: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 205–224). Waxmann.

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

Rost, D. H. (1991). Effect strength vs. statistical significance: A warning against the danger of small samples. *European Journal of High Ability*, *2*(2), 236–243. https://doi.org/10.1080/0937445910020212

Sadowski, I., & Zawistowska, A. (2020). The net effect of ability tilt in gendered STEM-related choices. *Intelligence*, *80*, Article 101439. https://doi.org/10.1016/j.intell.2020.101439

Sak, U. (2011). Prevalence of misconceptions, dogmas, and popular views about giftedness and intelligence: A case from Turkey. *High Ability Studies*, *22*(2), 179–197. https://doi.org/10.1080/13598139.2011.622942

Saß, S., Kampa, N., & Köller, O. (2017). The interplay of g and mathematical abilities in large-scale assessments across grades. *Intelligence*, *63*, 33–44. https://doi.org/10.1016/j.intell.2017.05.001

Schiepe-Tiska, A. (2019). *Mehrdimensionale Bildungsziele in internationalen Large-Scale Assessments: Konzeptualisierung, Entwicklung und Entstehungsbedingungen am Beispiel der MINT-Fächer* [Multidimensional educational goals in intenrational Large-Scale Assessments: Conceptualization, development and determining conditions in the context of STEM subjects] [Hablitation]. Technische Universität München, München. mediaTUM. https://mediatum.ub.tum.de/1523295

Schlatter, E., Molenaar, I., & Lazonder, A. W. (2021). Learning scientific reasoning: A latent transition analysis. *Learning and Individual Differences*, *92*, Article 102043. https://doi.org/10.1016/j.lindif.2021.102043

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

Simpson, R. L., LaCava, P. G., & Sampson Graner, P. (2004). The No Child Left Behind Act: Challenges and implications for educators. *Intervention in School and Clinic*, *40*(2), 67–75. https://doi.org/10.1177/10534512040400020101

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*, Article 2366. 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 promotion strategies for gifted adolescents]. *Reportpsychologie*, *29*(3), 170-176.

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K-12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research*, *86*(4), 849–899. https://doi.org/10.3102/0034654316675417

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

Steenbergen-Hu, S., Olszewski-Kubilius, P., & Calvert, E. (2020). The effectiveness of current interventions to reverse the underachievement of gifted students: Findings of a meta-analysis and systematic review. *Gifted Child Quarterly*, *64*(2), 132–165. https://doi.org/10.1177/0016986220908601

Sternberg, R. J., Chowkase, A., Parra-Martinez, F. A., & Landy, J. (2022). Criterion-referenced assessment of intelligence as adaptation to the environment: Is it possible, plausible, or practical? *Journal of Intelligence*, *10*(3), Article 57. https://doi.org/10.3390/jintelligence10030057

Stoeger, H. (2009). The history of giftedness research. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 17–38). Springer. https://doi.org/10.1007/978-1-4020-6162-2_2

Stoeger, H., Hopp, M., & Ziegler, A. (2017). Online mentoring as an extracurricular measure to encourage talented girls in STEM (science, technology, engineering, and mathematics): An empirical study of one-on-one versus group mentoring. *Gifted Child Quarterly*, *61*(3), 239–249. https://doi.org/10.1177/0016986217702215

Stoeger, H., Schirner, S., Laemmle, L., Obergriesser, S., Heilemann, M., & Ziegler, A. (2016). A contextual perspective on talented female participants and their development in extracurricular STEM programs. *Annals of the New York Academy of Sciences*, *1377*(1), 53–66. https://doi.org/10.1111/nyas.13116

Stumpf, E. (2012). *Förderung bei Hochbegabung* [Promotion for giftedness]. Kohlhammer.

Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, *7*(7), 935–947.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, *12*(1), 3–54. https://doi.org/10.1177/1529100611418056

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2019). High performance: The central psychological mechanism for talent development. In R. F. Subotnik, F. C. Worrell, & P. Olszewski-Kubilius (Eds.), *The psychology of high performance: Developing human potential into domain-specific talent* (pp. 7–20). American Psychological Association.

Taskinen, P. H., Schütte, K., & Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: The role of school factors, individual interest, and science self-concept. *Educational Research and Evaluation*, *19*(8), 717–733. https://doi.org/10.1080/13803611.2013.853620

Thomm, E., Sälzer, C., Prenzel, M., & Bauer, J. (2021). Predictors of teachers' appreciation of evidence-based practice and educational research findings. *Zeitschrift für Pädagogische Psychologie*, *35*(2-3), 173–184. https://doi.org/10.1024/1010-0652/a000301

Tieso, C. (2003). Ability grouping is not just tracking anymore. *Roeper Review*, *26*(1), 29–36. https://doi.org/10.1080/02783190309554236

Tirri, K., & Laine, S. (2017). Ethical challenges in inclusive education: The case of gifted students. In A. Gajewski (Ed.), *Ethics, equity, and inclusive education* (Vol. 9, pp. 239–257). Emerald Publishing. https://doi.org/10.1108/S1479-363620170000009010

Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Association for Supervision & Curriculum Development.

Trautmann, M., & Wischer, B. (2011). *Heterogenität in der Schule: Eine kritische Einführung* [Heterogeneity in schools: A critical introduction]. VS Verlag für Sozialwissenschaften.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, *98*(4), 788–806. https://doi.org/10.1037/0022-0663.98.4.788

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.

United Nations Educational, Scientific and Cultural Organization. (2009). *Policy guidelines on inclusion in education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000177849

United Nations Educational, Scientific and Cultural Organization. (2017). *A guide for ensuring inclusion and equity in education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000248254

Verband Bildung und Erziehung. (2016). *Multiprofessionelle Teams in der Schule* [Multi-professional teams at school]. https://www.vbe.de/fileadmin/user_upload/VBE/Themen/Positionen/2017_11_17_Position_Multiprofessionelle-Teams.pdf

Walsh, R. L., Kemp, C. R., Hodge, K. A., & Bowes, J. M. (2012). Searching for evidence-based practice: A review of the research on educational interventions for intellectually gifted children in the early childhood years. *Journal for the Education of the Gifted*, *35*(2), 103–128. https://doi.org/10.1177/0162353212440610

Weigand, G. (2019). Begabung und Begabungsförderung im Spiegel der Heterogenitätsdiskussion [Giftedness and the promotion of giftedness in the context of the heterogeneity discussion]. In C. Reintjes, I. Kunze, & E. Ossowski (Eds.), *Begabungsförderung und Professionalisierung: Befunde, Perspektiven, Herausforderungen* (pp. 73–85). Julius Klinkhardt.

Weigand, G., Fischer, C., Käpnick, F., Perleth, C., Preckel, F., Vock, M., & Wollersheim, H.-W. (Eds.). (2020). *Leistung macht Schule: Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler* [Promoting Excellence in School Education: Promotion of high-achieving and potentially high-achieving students]. Beltz.

Weis, M., Doroganova, A., Hahnel, C., Becker-Mrotzek, M., Lindauer, T., Artelt, C., & Reiss, K. (2019). Lesekompetenz in PISA 2018 - Ergebnisse in einer digitalen Welt [Reading literacy in PISA 2018 - Results in a digital world]. In K. Reiss, M. Weis, E. Klieme, & O. Köller (Eds.), *PISA 2018: Grundbildung im internationalen Vergleich* (pp. 47–80). Waxmann. https://doi.org/10.31244/9783830991007

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. https://www.nfer.ac.uk/publications/lgt01/lgt01.pdf

Wocken, H. (2019). *Inklusive Bildung: Annäherungen an den Begriff der Inklusion und Forderungen an die Inklusionsforschung* [Inclusive education: Approaches to the concept of inclusion and demands on inclusion research]. https://www.magazin-auswege.de/data/2019/09/Wocken_Inklusive_Bildung_Teil_1-4.pdf

Worrell, F. C. (2009). Myth 4: A single test score or indicator tells us all we need to know about giftedness. *Gifted Child Quarterly*, *53*(4), 242–244. https://doi.org/10.1177/0016986209346828

Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence*, *69*, 8–15. https://doi.org/10.1016/j.intell.2018.03.011

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

Ziegler, A., & Stoeger, H. (2012). Shortcomings of the IQ-based construct of underachievement. *Roeper Review*, *34*(2), 123–132. https://doi.org/10.1080/02783193.2012.660726

Ziernwald, L., Hillmayr, D., & Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms–A systematic review. *Journal of Advanced Academics*, *33*(4), 540–573. https://doi.org/10.1177/1932202X221112931

Ziernwald, L., Holzberger, D., Hillmayr, D., & Reiss, K. (2020). *Leistungsstarke Schülerinnen und Schüler fördern: Einblicke in Forschung und Praxis* [Encouraging high-achieving students: Insights into research and practice]. *Wissenschaft macht Schule*. Waxmann.

Ziernwald, L., Schiepe-Tiska, A., & Reiss, K. M. (2022). Identification and characterization of high-achieving student subgroups using two methodological approaches: The role of different achievement indicators and motivational-affective characteristics. *Learning and Individual Differences*, *100*, Article 102212. https://doi.org/10.1016/j.lindif.2022.102212

Appendix

Appendix A - Paper A

Ziernwald, L., Schiepe-Tiska, A., & Reiss, K. M. (2022). Identification and characterization of high-achieving student subgroups using two methodological approaches: The role of different achievement indicators and motivational-affective characteristics. *Learning and Individual Differences*, *100*, Article 102212. https://doi.org/10.1016/j.lindif.2022.102212

Learning and Individual Differences 100 (2022) 102212

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Identification and characterization of high-achieving student subgroups using two methodological approaches: The role of different achievement indicators and motivational-affective characteristics☆

Lisa Ziernwald*, Anja Schiepe-Tiska, Kristina M. Reiss

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

A R T I C L E I N F O

Keywords:
High-achieving
Identification
Cut-off score
Latent profile analysis (LPA)
Motivation

A B S T R A C T

How high-achieving student subgroups are identified is widely discussed. Studies use different domain-general and domain-specific achievement indicators and methodological approaches. Traditional research included cut-off scores, which have been criticized as arbitrary. Recently, latent profile analyses have been used more often. The present study compared these approaches regarding their overlap to investigate their interchangeability. Afterwards, high-achieving student subgroups were characterized by their motivational-affective characteristics. Data from $N = 1563$ high-achieving students were investigated. We used three achievement indicators (two mathematical competences and figural reasoning) for identification and four motivational-affective characteristics for characterization. The four largest high-achieving student subgroups were found through both approaches. However, the two methodological approaches could not be used interchangeably. Large heterogeneity in achievement indicators and motivational-affective characteristics existed across the subgroups. Top Performers have been identified as the target state for education, as they showed high scores on all achievement indicators and the most positive motivational-affective characteristics.

1. Introduction

One of the most dangerous and incorrect myths in gifted education is considering high-achieving students as one homogeneous group (Reis & Renzulli, 2009). High-achieving students are a many-faceted group with different potentials and competences in different achievement and motivational-affective areas. This heterogeneity is also reflected in multidimensional models describing high-achieving students. Multidimensionality is represented through (a) different achievement indicators and (b) that next to achievement indicators, motivational-affective characteristics are considered (Heller et al., 2005; Stoeger, 2009). Consequently, research has shifted its focus from investigating high-achieving students in general to examining high-achieving student subgroups specifically (Stoeger, 2009).

High-achieving students are identified using domain-general and domain-specific achievement indicators (Neuendorf et al., 2022). A study by Lohman et al. (2008) revealed that high-achieving students showed a larger intraindividual heterogeneity in achievement indicators compared to average-achieving students. A common distinction between achievement indicators is fluid versus crystallized intelligence. Fluid intelligence is defined as a “purely general ability to discriminate and perceive relations” and can thus be seen as domain-general achievement. Crystallized intelligence, in contrast, is defined as “discriminatory habits long established in a particular field” and can thus be seen as domain-specific achievement, such as mathematical competence (Cattell, 1943, p. 178). In the field of talent development, domain-general achievement indicators are often described as students' potential, and domain-specific achievement indicators are often described as manifest achievement in a certain domain (e.g., Heller et al., 2005; Preckel et al., 2020; Subotnik et al., 2011).

From a methodological perspective, traditionally, high-achieving student subgroups were identified using cut-off scores that separate a continuous variable into two categorical parts. These cut-off scores can easily be implemented in educational research and practice (e.g., de

☆ This research was funded by the Federal Ministry of Education and Research (BMBF) and the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK) [Grant number ZIB2016].

* Corresponding author at: Technical University of Munich, Centre for International Student Assessment (ZIB), Arcisstrasse 21, 80333 Munich, Germany.
E-mail addresses: lisa.ziernwald@tum.de (L. Ziernwald), schiepe-tiska@tum.de (A. Schiepe-Tiska), kristina.reiss@tum.de (K.M. Reiss).

https://doi.org/10.1016/j.lindif.2022.102212
Received 18 January 2022; Received in revised form 9 September 2022; Accepted 17 September 2022

Jonge et al., 2020; National Association for Gifted Children, 2010). However, from a research perspective, cut-off scores are criticized as relatively indiscriminate (de Jonge et al., 2020). Alternatively, latent profile analysis (LPA) was introduced as a person-centered approach that groups individuals into homogeneous subgroups (Hickendorff et al., 2018; Lazarsfeld & Henry, 1968; Mammadov et al., 2016). Both approaches have advantages and disadvantages (e.g., simply applicable vs. no arbitrary cut-off score needed) and are used in research to identify high-achieving student subgroups (e.g., Hofer & Stern, 2016; Leikin et al., 2017). However, thus far, only one methodological approach was applied simultaneously to identify high-achieving student subgroups from one student sample. Therefore, we do not know if those two often used methodological approaches lead to comparable and generalizable subgroups or how the choice of methodological approach affects the substantive results.

Further, multidimensional models describing high-achieving students not only refer to different achievement indicators but also consider motivational-affective characteristics (e.g., Heller et al., 2005). Together with achievement, motivational-affective characteristics form multidimensional educational goals, which are related to academic and occupational success (OECD, 2019; Schiepe-Tiska, 2019; Stoeger, 2009). Only when high-achieving students "are willing to take opportunities for growth in a domain and are able to persist when presented with challenges" ideal development of their potential can occur (Cross & Cross, 2017, p. 43). If teachers better understand the achievement and motivational heterogeneity of high-achieving students in their classroom, they could adapt their teaching strategies more specifically to the needs of different high-achieving student subgroups (Castejón et al., 2016).

The present study investigates the heterogeneity of high-achieving students by examining two research questions: (1) Do the cut-off score and the LPA approach result in comparable high-achieving student subgroups, and (2) How can these high-achieving student subgroups be characterized regarding their motivational-affective characteristics? To identify high-achieving student subgroups, we chose a domain-general and two domain-specific indicators. To investigate high-achieving students' heterogeneity in more detail, we characterized the different subgroups identified through achievement indicators through domain-general and domain-specific motivational-affective characteristics.

1.1. Identification of high-achieving students through achievement indicators

Various achievement indicators are used to identify high-achieving students (Carman, 2013; Neuendorf et al., 2022). Each resulting sample can be seen as one certain subgroup of the whole population of high-achieving students. A conceptual distinction exists between high-achieving students, often defined through high domain-specific academic achievement, and gifted students, often defined through high domain-general achievement. However, the identification methods used in empirical studies for high-achieving and gifted students showed a high overlap (Neuendorf et al., 2022). Therefore, we included in our definition of high-achieving students individuals with high domain-general (i.e., fluid intelligence) and/or high domain-specific (i.e., curriculum-oriented and application-oriented mathematical competence) achievement. Overall, our interest lay in high-achieving students capable of excellent achievement and who understand the social implications that, for example, mathematics plays in the world. This approach to high-achievement aligns with current global and political visions and movements that aim to enable a preferably broad student body to take on responsibility and leadership for the world's global challenges (Institute of Medicine, 2007; Kultusministerkonferenz [KMK], 2016a; OECD, 2020).

1.1.1. Fluid intelligence

Fluid intelligence considers decontextualized achievement without the need for content-specific knowledge. It is often measured through figural reasoning tasks (Baumert et al., 2009; Schulze et al., 2005), representing the ability to complete an evolving series of, for example, geometric figures (e.g., Wilhelm et al., 2014). Hence, fluid intelligence plays an important role in rapidly changing environments (Chiappe & MacDonald, 2005). On an aggregated level, fluid intelligence positively correlates with domain-specific achievements, such as mathematical competence (Mielicki et al., 2021; Saß et al., 2017; Taub et al., 2008). However, high fluid intelligence–often seen as students' potential–does not always translate into high achievement in a certain domain on an individual level (Gagné, 2004).

1.1.2. Curriculum-oriented and application-oriented mathematical competence

Neuendorf et al. (2022) found that high-achieving students are most often identified through tests. If one domain-specific indicator was used to operationalize high achievement, mathematics was by far the most common. Mathematical competence, which can be clearly distinguished from fluid intelligence (Baumert et al., 2009), is essential for lifelong learning and social participation (Ehmke et al., 2020). In large-scale assessments, two frameworks of mathematical competence, namely the application-oriented and the curriculum-oriented approach, are well known (Neidorf et al., 2006; Saß et al., 2017). Accordingly, the concept of mathematical competence can have two focuses: (a) the application of acquired knowledge in everyday life, as it has been operationalized in the Programme for International Student Assessment (PISA; de Lange, 2006; Neidorf et al., 2006; OECD, 2013a) and (b) curricular competences that are taught in classrooms, as it is operationalized in the National Assessment Study (NAS) or the National Assessment of Educational Progress (NAEP; Hartig & Frey, 2012; KMK, 2016b; Neidorf et al., 2006). There is an ongoing discussion if application-oriented and curriculum-oriented mathematical competences are independent concepts (Saß et al., 2017). On a theoretical level, the application-oriented and the curriculum-oriented frameworks both emphasize the special meaning of the context for applying knowledge. However, the application-oriented mathematical competence does not concentrate on curricular elements. It is defined as applying the knowledge learned at school in everyday events, namely in personal, occupational, societal, or scientific contexts, in order to answer real-world problems and realize the role of mathematics in the world (de Lange, 2006; Neidorf et al., 2006; OECD, 2013a). In contrast, in the curriculum-oriented framework, the contexts have a strong link to the national curriculum and the educational standards, and thus, competences are defined as skills that manifest themselves in a certain proficiency in special situations and focus on curricular validity (Pant et al., 2013a; Roppelt, Blum, & Pöhlmann, 2013).

Despite these different conceptual focuses on application- and curriculum-oriented mathematical competence, on an aggregated level, the assessed mathematical competences are highly correlated (r = 0.82–0.91; Ehmke et al., 2017; Klieme et al., 2001). Further, for grade 9 students, a comparable correlation between application-oriented (r = 0.70) and curriculum-oriented (r = 0.72) competence with fluid intelligence was found (Saß et al., 2017). However, on an individual level, students with an application-oriented mathematical competence score of 500 (PISA) reached scores between 380 and 610 in curriculum-oriented mathematical competence (NAS; Ehmke et al., 2017). Hence, these two constructs cannot be used interchangeably on an individual level. This distinction may also be necessary for identifying high-achieving student subgroups because differences for high-achieving students between these two competences were found in previous research (Lehner et al., 2017). Therefore, the current study investigates curriculum-oriented and application-oriented mathematical competence in the context of high-achieving students in more detail.

1.2. Methods of identifying high-achieving student subgroups

Two methodological approaches are often used to identify high-achieving student subgroups: cut-off scores and LPAs. The *cut-off score approach* dichotomizes a continuous variable into two parts and is often used in research and practice as it is easy to implement (e.g., de Jonge et al., 2020; National Association for Gifted Children, 2010). An example using cut-off scores is PISA, which ranks students on six competence levels and defines students on competence levels five and six as high-achieving (OECD, 2014a). Applying this definition to the domain of mathematics, 17.5 % of the students in Germany reached this cut-off score in PISA 2012 (OECD, 2014a, p. 298). In NAS, high-achieving students were defined as students scoring at the 95th percentile or higher (Roppelt, Penk, et al., 2013) or scoring at competence level 5 (Neuendorf et al., 2017). These examples show that different subgroups of students are identified as high-achieving depending on the used cut-off score. This heterogeneity in cut-off score use is in accordance with a systematic review that concluded that the cut-off scores used to identify high-achieving students varied considerably across studies. The median of the cut-off scores was a cut-off score of the top 10 % (Neuendorf et al., 2022). Two exemplary studies using cut-off scores on multiple achievement indicators to identify high-achieving student subgroups used an IQ score of 130 for general giftedness and two further cut-off scores on mathematics-related tests to define expertise in school mathematics. In a second step, differences between the detected subgroups were investigated (Leikin et al., 2014; Leikin et al., 2017). From these examples, it becomes clear that the identification of high-achieving students and high-achieving student subgroups is intermingled and that each resulting sample can be seen as one certain subgroup of the whole high-achieving student population. This rather arbitrary assignment of students into high-achieving student subgroups is a crucial point of criticism for using cut-off scores (Hickendorff et al., 2018; Stoeger, 2009).

One solution to bypass this criticism is using *LPAs*, which have been recently used more often in the field of high-achieving students (Hickendorff et al., 2018; Mammadov et al., 2016). LPAs are a suitable option to categorize persons into subgroups and investigate the relationship of these subgroups with other variables (Howard & Hoffman, 2018; Mammadov et al., 2016). LPAs explain the differences in a group of persons (e.g., high-achieving students) by defining homogeneous subgroups independent of a-priori-defined cut-off scores (Hickendorff et al., 2018). In order to gain homogeneous subgroups, the expectation-maximization (EM) algorithm is used. First, each subgroup's parameters (e.g., mean and variance) are guessed. The posterior probabilities based on these parameters are calculated in a second step. With those posterior probabilities, the parameters are re-estimated, and the posterior probabilities are re-calculated. This cycle is repeated until the parameters no longer change significantly (Oberski, 2016). The aim is to maximize "the homogeneity within subgroups while maximizing the heterogeneity between subgroups" (Mammadov et al., 2016, p. 177). Hofer and Stern (2016), for example, conducted LPAs with intellectual potential, operationalized through reasoning ability, and physics grades as achievement indicators to define five subgroups of students. They detected Overachievers ("very low intellectual potential […] and high physics grades"), High-achievers ("very high z-standardized intellectual potential […] and z-standardized physics grades"), Over-to-normal achievers ("rather low intellectual potential […] and average physics grades"), Under-to-normal achievers ("average intellectual potential […] and rather low physics grades"), and Underachievers ("high intellectual potential […] and very low physics grades"; Hofer & Stern, 2016, p. 125). One challenge for LPAs is whether the detected subgroups really exist or if they appeared because of methodological issues or the approach's exploratory character (Grimm et al., 2017).

To evaluate the comparability of the cut-off score and LPA approach on the aggregated level, it could be examined if the same subgroups can be found. Here, level differences, "the tendency for a person to be high, medium, or low across all factors", and shape differences, "the tendency for a person to have a distinct pattern of factors on which they are high, medium, or low" can be distinguished (Morin & Marsh, 2015, p. 39). On the individual level, one can test if both approaches would assign individual students to the same high-achieving subgroups. If both approaches result on an aggregated level in the same subgroups and assign the same students on an individual level to each subgroup, both methodological approaches could be used interchangeably.

1.3. Multidimensional educational goals and the characterization of high-achieving students

Besides achievement indicators, motivational-affective characteristics are important for characterizing high-achieving students (Subotnik et al., 2011). They may explain why or why not students' potential (i.e., domain-general achievement) transfers into domain-specific manifest achievement (e.g., mathematical competence; Cross & Cross, 2017; Heller et al., 2005) and are (reciprocally) related to domain-specific achievement indicators (e.g., Arens et al., 2017; Lee & Stankov, 2018; OECD, 2013c). Like achievement indicators, motivational-affective indicators can be divided into domain-general versus domain-specific characteristics (Kuger & Klieme, 2016). The present study focused on perseverance as domain-general and mathematics self-efficacy, interest, and instrumental motivation as domain-specific characteristics, which are particularly important in the context of high-achieving students. They correlate quite highly with academic achievement and are investigated in previous research with high-achieving students (e.g., OECD, 2013c).

Perseverance describes the willingness to persist on difficult problems without giving up, even if barriers appear (OECD, 2013c). The willingness to persist is an important aspect of talent development as it can be seen as a prerequisite that potential transfers into manifest achievement (Cross & Cross, 2017). Research showed that students who, on average, show more perseverance in problem-solving achieve higher mathematics achievement (Cutumisu & Bulut, 2017; Fung et al., 2018).

Self-efficacy regarding mathematics is "an individual's confidence in his or her ability to perform mathematics and is thought to directly impact the choice to engage in, expend effort on, and persist in pursuing mathematics" (Ashcraft & Rudig, 2012, p. 249). This domain-specific definition is in accordance with Bandura's definition of self-efficacy, defined as "beliefs in one's capabilities to organize and execute the courses of action required to manage prospective situations. Efficacy beliefs influence how people think, feel, motivate themselves, and act" (Bandura, 1995, p. 2). When students have doubts about mastering a mathematics task, they may not try to perform at all (Ritchotte et al., 2014). In PISA, mathematics self-efficacy was the strongest predictor of mathematics achievement compared to 44 other investigated motivational-affective and socio-economic status-related predictors (Lee & Stankov, 2018), even after controlling for other motivational-affective characteristics (Pitsia et al., 2017).

Mathematics *interest* represents the affective component of a relationship between a person and an object (Krapp & Prenzel, 2011; OECD, 2014b) and corresponds to the concept of intrinsic motivation, according to Ryan and Deci (2000). Learning activities based on interest are done for their own sake and not for external rewards (OECD, 2013a; Ryan & Deci, 2000). Interest is required to voluntarily engage in certain activities, such as choosing a certain field of study (Taskinen et al., 2013). This aspect is of special importance as there is a high need for graduates in mathematics (Cappelli, 2015; Institute of Medicine, 2007), and "the level of a person's interest has repeatedly been found to be a powerful influence on learning" (Hidi & Renninger, 2006, p. 111). Mathematics interest and achievement are positively related. Thus, students more interested in mathematics also show higher achievement (Fung et al., 2018; Lee & Stankov, 2018).

Instrumental motivation regarding mathematics refers to reaching a specific goal beyond completing the task itself, such as striving for a job in the field of mathematics. This type of motivation focuses on the instrumental value of the task and can foster students' motivation for a certain domain and corresponds to the concept of extrinsic motivation

(Lens et al., 2009; OECD, 2013a; Ryan & Deci, 2000). In PISA 2012, most students across OECD countries agreed that learning mathematics would enhance their job perspectives (OECD, 2013c). Moreover, instrumental motivation and mathematics achievement were positively related (OECD, 2013c; Pitsia et al., 2017).

Studies investigating differences in these motivational-affective characteristics between high-achieving and not-as-high-achieving identified students mainly revealed more positive characteristics for high-achieving students (Agaliotis & Kalyva, 2019; Gubbels et al., 2018; Hong & Aqui, 2004; Litster & Roberts, 2011). Further, PISA 2012 showed that the relationship between these motivational-affective characteristics and mathematics achievement is, on average, higher for high-achieving students compared to lower-achieving students (OECD, 2013c). One subgroup of high-achieving students, which has been investigated in more detail concerning its motivational-affective characteristics, are the Underachievers. They showed lower intrinsic motivation, motivational orientations, interest, self-efficacy, and effort than high-achieving peers who fulfill their potential (Agaliotis & Kalyva, 2019; Hofer & Stern, 2016; Lüftenegger et al., 2015; McCoach & Siegle, 2003). Other high-achieving student subgroups were investigated less frequently (e.g., Castejón et al., 2016; Gubbels et al., 2018). Hong and Aqui (2004) summarized that–despite the advancement in conceptualizing high achievement as multidimensional–the comparison of different high-achieving student subgroups regarding their cognitive and motivational-affective characteristics is not sufficient yet.

2. Present study

We decided to define students scoring on at least one achievement indicator in the top 20 % as high-achieving, as we aimed to investigate next to the already top-performing students also the potentially high-achieving students (Friedman & Lee, 1996). This broader perspective also follows Schmidtner (2017), who argued that the top 20 % of a student population is a useful definition of high-achieving students as this could be seen as a realistic international threshold for high-achieving students in a country. Further, this student population is defined as students needing regular curriculum modifications, enrichment, and differentiation in the educational context (Renzulli, 1984). Therefore, a clear understanding of the heterogeneity of these high-achieving students is necessary.

The present study expands previous research by applying the cut-off score *and* the LPA approach to the same high-achieving student sample to investigate whether both approaches result in comparable high-achieving student subgroups (RQ1). We used three achievement indicators that focus on different aspects of achievement: fluid intelligence, operationalized through figural reasoning (Wilhelm et al., 2014), curriculum-based mathematical competence, operationalized through mathematical competence in Germany's NAS (Pant et al., 2013a), and application-oriented mathematical competence, operationalized through mathematical competence in PISA (de Lange, 2006; OECD, 2013a). A comprehensive overview, including the theoretical background and crucial characteristics of the used achievement indicators, can be found in Appendix A.

The overlap between the cut-off score and the LPA approach was examined on an aggregated and individual level. We, therefore, considered shape and level differences, the sample size of the detected subgroups, and the assignment of each individual to the high-achieving student subgroups. As certain high-achieving student subgroups (i.e., Overachievers, Underachievers, and Top Performers) were detected in previous studies through the cut-off score (e.g., Agaliotis & Kalyva, 2019; Leikin et al., 2017) and the LPA approach (e.g., Hofer & Stern, 2016), we expected to find a comparable set (e.g., same number, same size, and the same level and shape differences) of high-achieving student subgroups across both methodological approaches. However, to our knowledge, thus far, no study has applied both approaches to the same sample.

Further, we examined how the identified high-achieving student subgroups differed in their motivational-affective characteristics (RQ2). Little research has examined motivational-affective characteristics in different high-achieving student subgroups, although their importance has been theoretically and empirically confirmed (Heller et al., 2005; OECD, 2013c; Stoeger, 2009). We used the domain-general characteristic perseverance and mathematics-related characteristics self-efficacy, interest, and instrumental motivation. Based on theories and previous research that established a positive relationship between achievement and motivational-affective characteristics (e.g., Bandura, 1977; Krapp & Prenzel, 2011; Lee & Stankov, 2018; OECD, 2013c), we expected that high-achieving students who fulfill their potential in all investigated achievement indicators (i.e., Top Performers) show more positive motivational-affective characteristics compared to other subgroups of high-achieving students (hypothesis 2a). Furthermore, in line with previous research (e.g., Agaliotis & Kalyva, 2019; Hofer & Stern, 2016), we expected that high-achieving students who do not fulfill their potential (i.e., Underachievers) show worse motivational-affective characteristics compared to performing high-achievers (hypothesis 2b). We did not specify further hypotheses for other high-achieving student subgroups due to the lack of previous research.

3. Method

3.1. Sample and procedure

Data from a German national expansion of PISA 2012—a class-based sample of two ninth grades from each participating school—were used to unite two special features of this PISA dataset. First, mathematical competence was the main domain, and second, data could be linked to Germany's NAS 2012. On the first day of data collection, the two-hour cognitive PISA assessment took place, and the student background questionnaire, including the motivational-affective characteristics, was administered with a rotated booklet design. On the second day, students participated in the assessment of NAS mathematical competence and figural reasoning. We considered all ninth-graders with values on all three achievement indicators ($N = 5154$). From this sample, for the definition of our high-achieving student sample, we used a liberal cut-off score and defined the top 20 % of students on at least one of the three achievement indicators (or-approach) as high-achieving (cf. Renzulli, 1984; Schmidtner, 2017). Students who reached the 80th percentile in none of the achievement indicators were defined as non-high-achieving and not further analyzed. Our final sample consisted of $N = 1563$ high-achieving students ($M_{age} = 15.30$ years, $SD = 0.46$, 45.6 % female) originating from 180 different schools and 308 school classes. 78.7 % of our high-achieving sample attended the highest track of the German secondary school system (Gymnasium).

As data collection took place in the context of PISA and NAS, it was commissioned by the federal governments of Germany. According to school laws, participation in the competence test was mandatory. Participation in the background questionnaire was only mandatory in some federal states. If participation was not mandatory, parents and students gave written informed consent. The current study was conducted according to the APA ethical standards. A review through an ethics committee was not necessary by institutional guidelines or national regulations, following the German Research Foundation, as the used data were anonymized, and no further disclosure is possible (cf. Schiepe-Tiska, 2019; Siegle et al., 2013).

3.2. Measures

All measures used in the current study are established research tools, which have been tested regarding their quality criteria through the PISA and NAS committees (Lenski et al., 2016; OECD, 2013a). The descriptive statistics and intercorrelations of the used measures are displayed in Table 1.

Table 1
Descriptive statistics and Pearson intercorrelations for achievement indicators and motivational-affective characteristics ($N = 1563$).

		Mean	SD	1	2	3	4	5	6
1	EAP PISA	0.99	0.64	–					
2	EAP NAS	1.04	0.65	0.66*	–				
3	EAP BEFKI	0.79	0.70	−0.04	0.11*	–			
4	PERSEV	0.10	0.96	0.20*	0.17*	0.04	–		
5	MATHEFF	0.71	0.93	0.42*	0.36*	0.03	0.39*	–	
6	INTMAT	−0.02	1.05	0.19*	0.20*	0.15*	0.43*	0.37*	–
7	INSTMOT	−0.06	1.01	0.17*	0.16*	0.10*	0.37*	0.33*	0.63*

Note. EAP = expected a posteriori estimator; PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning; PERSEV = perseverance; MATHEFF = mathematics self-efficacy; INTMAT = mathematics interest; INSTMOT = instrumental motivation for mathematics.
* $p < .05$.

3.2.1. Achievement indicators

Fluid intelligence was measured through figural reasoning tasks, representing an adequate and often used possibility to measure fluid intelligence (Baumert et al., 2009; Schulze et al., 2005). An unpublished parallel form of the published subscale *deductive reasoning* of a test for fluid and crystallized intelligence (Berliner Test zur Erfassung der Fluiden und Kristallinen Intelligenz [BEFKI]; Wilhelm et al., 2014) was used. The deductive reasoning subscale consisted of 16 tasks in which students were asked to complete a row of three geometric figures. Correct answers were scored with *1*, and incorrect answers were scored with *0*.

Application-oriented mathematical competence was measured by the mathematics global scale in PISA, which contained four content areas: quantity, uncertainty and data, change and relationships, and space and shape (OECD, 2014a, 2014b). Sample items can be found in OECD (2014a) and OECD (2013b). For each student, five plausible values were estimated. The students' test scores can be assigned to seven proficiency levels, ranging from below level 1 to level 6 (see OECD, 2014b, also for a detailed description of the proficiency levels).

Curriculum-oriented mathematical competence was assessed by the mathematics global scale in NAS, which contained five content areas: numbers, measurement, data and chance, functional relationships, and space and shape. All items had a strong link to the national curriculum and the educational standards in Germany (Pant et al., 2013a; Roppelt, Blum, & Pöhlmann, 2013). For each student, 15 plausible values were estimated (Hecht et al., 2013). Students' test scores can be assigned to six proficiency levels, ranging from levels 1a and 1b to level 5 (see Pant et al., 2013b, also for a detailed description of the proficiency levels). One example item of application-oriented (PISA) and one of curriculum-oriented (NAS) mathematical competence, including a comparison of both, can be found in Lehner et al. (2017). Appendix B shows two released sample items for mathematical competence from PISA and NAS that further illustrate the different foci on application-oriented versus curriculum-oriented mathematical competence.

3.2.2. Motivational-affective characteristics

Perseverance (PERSEV) was assessed with five items (e.g., "I continue working on tasks until everything is perfect"), which were answered on a five-point Likert scale (1 = *very much like me* to 5 = *not at all like me*). Three items were reversed so that higher scores indicate higher perseverance (OECD, 2014b, p. 337). The scale's reliability is α = 0.75 (N = 950).

Mathematics Self-Efficacy (MATHEFF) was assessed with eight items. Students rated how confident they feel in handling specific mathematical tasks (e.g., "Calculating how much cheaper a TV would be after a 30% discount") on a four-point Likert scale (1 = *very confident* to 4 = *not at all confident;* OECD, 2014b*, p. 322*). The scale's reliability is α = 0.78 (N = 945).

Mathematics Interest (INTMAT) was assessed with four items (e.g., "I enjoy reading about mathematics"), which were answered on a four-point Likert scale (1 = *strongly agree* to 4 = *strongly disagree;* OECD, 2014b*, p. 321*). The scale's reliability is α = 0.89 (N = 947).

Instrumental Motivation for Mathematics (INSTMOT) was assessed with four items (e.g., "I will learn many things in mathematics that will help me get a job"), which were also answered on a four-point Likert scale (1 = *strongly agree* to 4 = *strongly disagree;* OECD, 2014b, p. 322). The scale's reliability is α = 0.87 (N = 950). All domain-specific motivational-affective items were reversed so that higher scores indicate higher motivational-affective characteristics (OECD, 2014b, pp. 321–322).

3.3. Statistical analyses[1]

First, high-achieving student subgroups were identified through the cut-off score and the LPA approach using three achievement indicators. We used the plausible values (PVs) of each achievement indicator from the PISA and NAS datasets that can be defined "as random values from the posterior distributions." (OECD, 2009, p. 95). To ensure comparability of the achievement indicators, we z-standardized the achievement indicators. Therefore, we calculated one vector of all PVs (5 for PISA and 15 for NAS and BEFKI) for each achievement indicator as they originate from one common distribution and determined the means and standard deviations for each vector for the whole student sample (N = 5154) using the software R (Version R 4.1.3). These values were used for z-standardization so that 0 represents the average value of all ninth-grade students (high-achieving and not-as-high-achieving identified) in our sample, and 1 represents the standard deviation for each achievement indicator. Next, the expected a posteriori estimators (EAPs) were used for all achievement indicators. The EAP represented the posterior distribution's average and could be used as "an individual estimate of student ability" (OECD, 2009, p. 97; Rost, 2004). Therefore, we calculated the EAP for each achievement indicator separately by averaging the respective standardized PVs.

For the *cut-off score approach* for each achievement indicator, the 80th percentile of each plausible value was calculated and averaged afterwards for the whole student sample (N = 5154). Next, students were dichotomously categorized as above or below the 80th percentile on each indicator. The result were seven high-achieving student subgroups (Table 2).

The names of the subgroups were assigned according to the theoretical conceptions and main characteristics of the three achievement indicators (Appendix A). Students who did not reach the 80th percentile on any of the three achievement indicators were defined as non-high-achieving and not considered for further analyses. For all analyses in this step, SPSS Version 28 was used.

For the *LPA approach*, we conducted five sets of LPAs with different variance-covariance matrix specifications and different numbers of profiles (cf. Pastor et al., 2007). Each model was estimated with 5000 random selections of starting values, 500 iterations for each random

[1] The syntax files for the analysis steps of our research questions can be accessed via an OSF project. Please contact the corresponding author of the manuscript to gain access.

Table 2
Description of the defined high-achieving student subgroups.

Subgroup	Top 20 % PISA = application-oriented mathematics	Top 20 % NAS = curriculum-oriented mathematics	Top 20 % BEFKI = figural reasoning	Verbal description Students show …
Top Performer	Yes	Yes	Yes	High application- and curriculum-oriented mathematical achievement as well as high figural reasoning achievement.
Overachiever	Yes	Yes	No	High application- and curriculum-oriented mathematical achievement. Their figural reasoning achievement is average.
Wise Learner	No	Yes	Yes	High curriculum-oriented mathematical and high figural reasoning achievement. Their application-oriented mathematical achievement is average.
Wise Operator	Yes	No	Yes	High application-oriented mathematical and high figural reasoning achievement. Their curriculum-oriented mathematical achievement is average.
Operator	Yes	No	No	High application-oriented mathematical achievement. Their curriculum-oriented mathematical and their figural reasoning achievement are average.
Learner	No	Yes	No	High curriculum-oriented mathematical achievement. Their application-oriented mathematical and their figural reasoning achievement are average.
Underachiever[a]	No	No	Yes	High figural reasoning achievement. Their application- and curriculum-oriented mathematical achievement are average.

[a] The two Underachiever subgroups (1 and 2) defined by the latent profile analyses differed regarding their level in the achievement indicators. The qualitative characteristics described in this table refer to all Underachiever subgroups detected in this manuscript. The subgroup of students who reached the 80th percentile in none of the achievement indicators was defined as non-high-achieving and was not further investigated.

start, and the 500 best solutions for final optimization. To find the appropriate number of latent profiles, we compared statistical indicators of the two to nine profile solutions: Akaike information criteria (AIC), Bayesian information criteria (BIC), sample-size adjusted BIC (SABIC), entropy, and adjusted Vuong-Lo-Mendell-Rubin likelihood ratio test (aVLMRT). Lower values on the information criteria indicate a better model fit. The aVLMRT test is not significant when the model with k-1 profiles shows a better fit than the model with k profiles. The entropy ranges from 0 to 1, and a higher value indicates a more accurate classification of individuals (Geiser, 2013; Hong et al., 2020; Nylund et al., 2007). However, as none of these statistical indicators can be seen as the gold standard for model selection (Mammadov et al., 2016), we also considered the content-related interpretability, the parsimony, and the number of persons per profile (cf. Berlin et al., 2014). One useful feature of LPA is that classification probabilities are modeled (Bacher et al., 2010). This advantage could not be used in the current study, as we used imputed data sets to analyze our motivational-affective outcomes. For analyses with type imputation, the three-step approach in MPlus is not available. Thus, each student was assigned "to the latent profile with the highest posterior membership probability" (Mammadov et al., 2016, pp. 180–181). Consequently, we did not use the classification probabilities for further analyses and could not consider the classification uncertainties for parameter estimation. In line with the cut-off score approach, the subgroup names were given according to the three achievement indicators' main characteristics (Appendix A). LPAs were conducted with the software Mplus (Version 8). For the automatized execution of Mplus syntax, we used the R package MplusAutomation (R Version R 4.1.3; Hallquist & Wiley, 2018). The achievement indicator values for each subgroup were calculated using SPSS Version 28.

For the motivational-affective characteristics, two types of missing values were observed because of the questionnaire's rotated booklet design: missing values by design (i.e., missing completely at random; Enders, 2010) and missing values because students did not answer each item (Table 3). To handle missing values, multiple imputation was conducted for all ($N = 5154$) students (Enders, 2010). The common part of the student background questionnaire was used for the imputation model. These variables contained information about students' study program, age, gender, preschool education, age when starting school, grade repetition, truancy, family structure, immigrant background,

Table 3
Missing values of the motivational-affective indicators.

	PERSEV	MATHEFF	INTMAT	INSTMOT
Valid cases	957	960	960	958
MCAR	532 (87.79)	532 (88.23)	532 (88.23)	532 (87.93)
Other missing values	74 (12.21)	71 (11.77)	71 (11.77)	73 (12.07)

Note. The percentage of missing values due to the booklet design (missing completely at random; MCAR) and missing values due to unknown reasons are in brackets, respectively. PERSEV = perseverance; MATHEFF = mathematics self-efficacy; INTMAT = mathematics interest; INSTMOT = instrumental motivation for mathematics.

language spoken at home, ESCS, and the number of books at home. The technical report gives a full overview of the common part, including item numbers (OECD, 2014b). Three items of the common part were not used for the imputation model as they had the same manifestation for all students (i.e., grade), missings for all students (i.e., grade retention on ISCED Level 3), or less than 5 % valid scores (i.e., age of arrival in test country). Further, the detected student subgroups were included as dummy variables in the imputation model as these variables were used in the analysis phase. Fifty imputed data sets were created (cf. Graham et al., 2007).

In the analysis phase, for each of the 50 data sets and each motivational-affective characteristic, a regression analysis with subgroups as dichotomous independent variables (0 = *not belonging to the respective subgroup* and 1 = *belonging to the respective subgroup*) and the respective motivational-affective characteristic as the dependent variable was conducted. Top Performers represented the reference group because they scored highly on all three achievement indicators. Hence, this subgroup is considered as the target state. Bonferroni-Holm correction was applied within each regression analysis to account for multiple comparisons with the reference group (six comparisons for the cut-off score subgroups and four for the LPA subgroups). All indices of the motivational-affective characteristics were standardized “with an OECD average of 0 and a standard deviation of 1” (OECD, 2014b, p. 314). Mplus (Version 8) was used to perform these analyses.

4. Results

4.1. Preliminary analyses

As the original dataset used a class-based sampling technique, we investigated whether the allocation to different subgroups was independent of the nested structure. In the cut-off score subgroups the average cluster size ranged between $M_{min} = 1.16$ and $M_{max} = 2.51$ ($Mdn_{min} = 1$, $Mdn_{max} = 2$) and in the LPA subgroups between $M_{min} = 1.23$ and $M_{max} = 3.39$ ($Mdn_{min} = 1$, $Mdn_{max} = 3$). From these small clusters within each subgroup, it can be assumed that the allocation to the different subgroups was almost completely independent of the students' clustering within school classes.

Table 1 presents descriptive statistics and intercorrelations of the achievement indicators and motivational-affective characteristics for the high-achieving student sample. The means and standard deviations of the achievement indicators for each detected subgroup separately can be found in Table 4.

The correlations of the achievement indicators within each subgroup can be found in Appendix C. Appendix D presents sample characteristics (i.e., percentage of girls and students attending the highest secondary school track) for each detected subgroup separately.

4.2. High-achieving student subgroups – aggregated level

Fig. 1 presents the means and standard errors of the achievement indicators as well as the number of students for each of the seven subgroups detected through the cut-off score approach.

Table 4
Descriptive statistics of the achievement indicators for each subgroup.

	EAP_PISA		EAP_LV		EAP_NAS	
	M	*SD*	*M*	*SD*	*M*	*SD*
Top Performer	1.50	0.41	1.62	0.42	1.36	0.42
Overachiever	1.37	0.36	1.38	0.35	0.28	0.46
Underachiever	0.20	0.51	0.26	0.45	1.20	0.33
Operator	1.13	0.23	0.47	0.36	0.06	0.52
Learner	0.53	0.28	1.12	0.21	0.17	0.50
Wise Learner	0.55	0.28	1.19	0.30	1.29	0.40
Wise Operator	1.16	0.22	0.69	0.19	1.22	0.25
Top Performer	1.66	0.40	1.83	0.31	1.15	0.59
Overachiever	0.97	0.41	1.11	0.24	0.63	0.70
Underachiever 1	−0.10	0.49	−0.03	0.41	1.21	0.31
Underachiever 2	0.58	0.33	0.59	0.17	1.07	0.13
Operator	1.09	0.17	0.39	0.43	−0.03	0.56

Note. EAP = expected a posteriori estimator; PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning. The upper part of the table represents the cut-off score subgroups, and the lower part the LPA subgroups.

The Top Performers formed the largest subgroup, with scores at least at the 80th percentile in all three achievement indicators (cf. Table 2 for a description of the subgroups). The Overachievers formed the second largest subgroup and scored at or above the 80th percentile in both mathematical competences but not in figural reasoning. One-fifth of the students were classified as Underachievers. These students exceeded the cut-off score in figural reasoning but not in mathematical competences. An already noticeable smaller group of students was classified as Operators, who showed high achievement in application-oriented mathematics, but not in curriculum-oriented mathematics and figural reasoning. Their group size was comparable to the Learners, who, in contrast, achieved scores at or above the 80th percentile in curriculum-oriented mathematics but not in application-oriented mathematics and figural reasoning. The Wise Learners formed the second smallest subgroup and exceeded the cut-off score in figural reasoning and curriculum-oriented mathematics but not in application-oriented mathematics. The Wise Operators formed the smallest subgroup. These students scored at or above the 80th percentile in figural reasoning and application-oriented mathematics but not in curriculum-oriented mathematics.

Table 5 shows the model fit criteria for the class-varying diagonal model of the LPAs and the number of students in each subgroup. The AIC and SABIC consistently declined with increasing numbers of profiles and, therefore, did not help to identify the profile solution that fit the data best. The BIC declined until the eight profile solution. The BIC differences between profile solutions showed that the decline after the five profile solution got noticeably smaller. The aVLMRT was not significant for the seven profile solution and thus indicated a better fit of the six profile solution than the seven profile solution. When we compared the five, six, and seven profile solutions on the content level, we found that one subgroup, which disappeared from the six and seven profile solution to the five profile solution, was very similar to the still existing Underachiever 1 subgroup. As these two subgroups were rather small (approximately 5 % and 6 %, cf. Appendix E) and showed a comparable pattern of achievement indicators, we decided to use the five profile solution, which combined these two subgroups into one. The second subgroup that disappeared from the seven to the five profile solution was very similar to two still existing subgroups (Top Performers and Overachievers) and lay in between them (Appendix E).

Hence, we selected the more parsimonious solution with five subgroups, which had a similar model fit as the six and seven profile solution. Thus, the LPA found an Operator subgroup, two Underachiever subgroups, a Top Performer subgroup, and an Overachiever subgroup. Fig. 2 presents for each of the five LPA subgroups the means and standard errors of the achievement indicators as well as the number of students in each subgroup. Overall, four subgroups defined through the

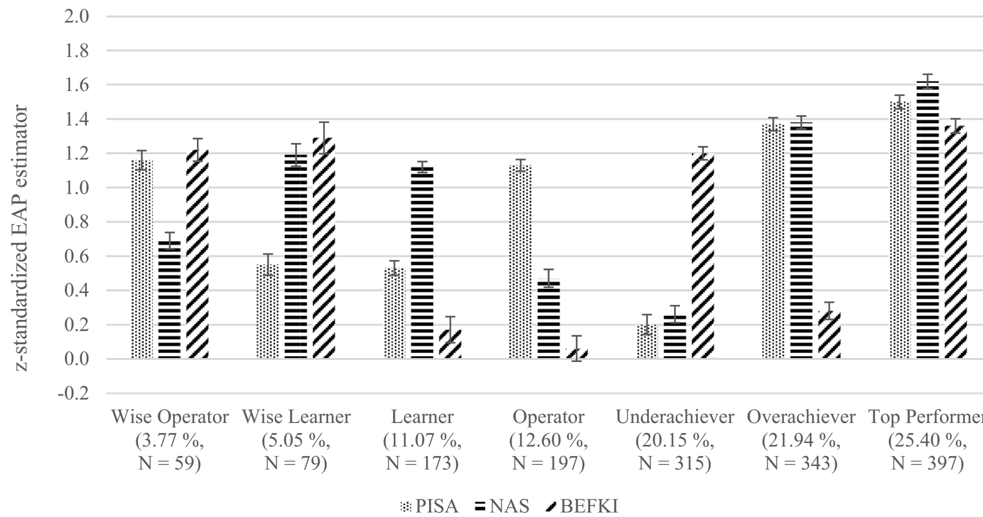


Fig. 1. Means (±2 SE) of the achievement indicators for each cut-off score subgroup.
Note. PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning. The subgroups are arranged according to their size in ascending order.

Table 5
Fit results from the latent profile analyses with free variances across profiles and no covariances.

Profiles	LL	par	AIC	BIC	SABIC	Entropy	aVLMRT	BIC diff. value	Number of students in profiles[a]								
									1	2	3	4	5	6	7	8	9
2	−4368.917	13	8763.83	8833.44	8792.14	0.850	$p < .001$	x	238	1325							
3	−4102.218	20	8244.44	8351.52	8287.99	0.788	$p < .001$	481.92	1016	208	339						
4	−4032.573	27	8119.15	8263.71	8177.94	0.722	$p = .024$	87.81	842	113	239	369					
5	−3982.828	34	8033.66	8215.70	8107.69	0.730	$p = .024$	48.01	151	168	140	365	739				
6	−3943.783	41	7969.57	8189.09	8058.85	0.744	$p = .047$	26.61	75	99	155	139	729	366			
7	−3910.679	48	7917.36	8174.37	8021.88	0.711	$p = .139$	14.73	138	571	161	77	99	353	164		
8	−3876.739	55	7863.48	8157.97	7983.25	0.736	$p = .337$	16.40	262	83	174	18	110	136	157	623	
9	−3853.755	62	7831.51	8163.48	7966.52	0.753	$p = .050$	−5.51	80	111	20	136	151	625	137	286	17

Note. LL = model log-likelihood; par = number of estimated parameters; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; SABIC = sample-size adjusted BIC; aVLMRT = Adjusted Vuong-Lo-Mendell-Rubin likelihood ratio test; BIC diff. value = difference between the BIC values of profile n-1 and n.
[a] Depending on the assumed number of profiles, students assigned to a specific profile are not comparable across profile solutions.

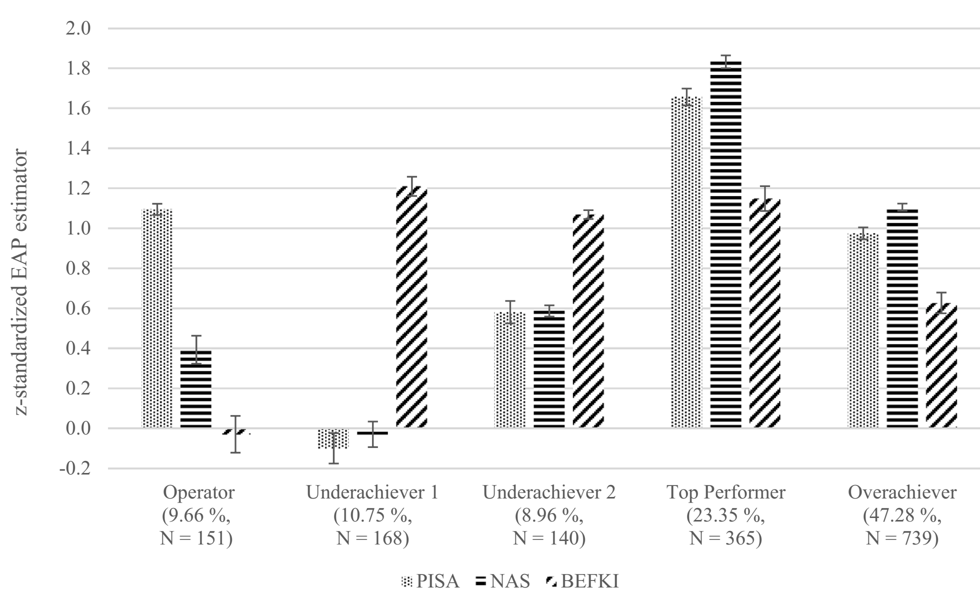


Fig. 2. Means (±2 SE) of the achievement indicators for each LPA subgroup.
Note. PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning.

cut-off score approach, namely Top Performers, Overachievers, Underachievers, and Operators, could also be found using LPAs.

4.3. Comparison of the two approaches – individual level

In order to test whether both approaches classified the same high-achieving students into each subgroup, a cross table between the subgroups was created (Table 6).

Students identified as Underachievers by the cut-off score approach were almost solely represented in the two corresponding LPA Underachiever subgroups. The same was true for the majority of the Operators. They also were mainly represented in the corresponding LPA Operator subgroup. The Overachievers and Top Performers detected through the cut-off score approach were predominantly comprised of the respective LPA subgroup. However, a quite high overlap between these two subgroups was also detected. The Learners were completely merged into the LPA Overachiever subgroup, which was also mostly true for the Wise Learners and Wise Operators. Comparing the two methodological approaches on an individual level yielded that most students (59 %) were grouped into the same subgroup independent of the methodological approach used.[2] The relationship between the two methodological approaches was Cramer's $V = 0.67$, which can be interpreted as a strong relationship (Cohen, 1977).

We further conducted sensitivity analyses to determine whether a different criterion for defining the high-achieving student sample and subgroups would have changed our results. Hence, we used high-achieving students as the overall analysis sample who scored at or above the 90th percentile in at least one of the three achievement indicators ($N = 839$). With this subsample, we again conducted the cut-off score and LPA approach with the class-varying diagonal model. The detected subgroups from these analyses showed very similar patterns compared to our main analyses (see Appendices F and G). Although the overlap between the two approaches on an individual level was lower (47 %) when the 90th percentile was used as the cut-off score, on an aggregated level, comparable results regarding high-achieving student subgroups across the different cut-off scores were found.[3]

4.4. Motivational-affective characteristics of the high-achieving student subgroups

Table 7 presents the descriptive statistics for each detected subgroup's imputed motivational-affective characteristics. Table 8 presents the regression analyses of the motivational-affective characteristics for the cut-off score subgroups. All subgroups consistently showed significantly lower motivational-affective characteristics than the Top Performers. One exception is the smallest subgroup containing the Wise Operators, who showed no significant differences from Top Performers' perseverance, mathematics interest, and instrumental motivation for mathematics. Second, Wise Learners showed no significant differences in instrumental motivation for mathematics from Top Performers. Overall, next to the Top Performers, Wise Operators and Overachievers showed relatively high motivational-affective characteristics.

Table 9 presents the regression analyses of the motivational-affective characteristics for the LPA subgroups, which were similar to the cut-off score subgroups. Again, Top Performers showed the highest perseverance, mathematics self-efficacy, mathematics interest, and instrumental motivation for mathematics. All other subgroups showed significantly lower motivational-affective characteristics.

Independent of the used approach, the four subgroups found by both methodological approaches (i.e., Top Performers, Overachievers, Operators, and Underachievers) showed similar results regarding their motivational-affective characteristics. Top Performers showed the highest motivational-affective characteristics. Overachievers, Operators, and Underachievers showed significantly lower scores across all motivational-affective characteristics.

5. Discussion

The present study examined the identification of high-achieving student subgroups by comparing the cut-off score and LPA approach (RQ1) and characterized the identified subgroups regarding their motivational-affective characteristics (RQ2). Key findings were: (1) On the aggregated level, both approaches revealed the four largest high-achieving student subgroups. On the individual level, there was an overlap of 59 %. (2) Top Performers showed the highest motivational-affective characteristics, whereas Underachievers showed relatively low motivational-affective characteristics independent of the methodological approach used.

5.1. High-achieving student subgroups based on achievement indicators

The cut-off score and the LPA approach only partially identified comparable high-achieving student subgroups. Both approaches revealed the following subgroups that have also been identified in previous research investigating subgroups of high-achieving students using multiple achievement indicators (e.g., Agaliotis & Kalyva, 2019; Hofer & Stern, 2016; Leikin et al., 2017): Top Performers, Overachievers, and Underachievers. In addition, a group of Operators–defined through high application-oriented mathematical competence–was found through both approaches. Although different numbers of subgroups emerged from the methodological approaches, the overlap of assigning students to the same subgroups (59 %) can be rated as quite high. In particular, both approaches identify the larger–and thus, for educational practice, most relevant–subgroups of high-achieving students quite reliable.

The Learners, Wise Learners, and Wise Operators detected through the cut-off score approach were mostly assigned to the LPA Overachievers. The result that the LPA Overachiever subgroup is scattered across all subgroups identified by the cut-off score approach could be explained methodologically. As the LPA Overachievers show achievement scores quite close to the achievement scores of the overall high-achieving student sample and the cut-off scores of the three achievement indicators, the distance between the LPA Overachievers and students from other subgroups might be quite small. As the LPA aims to maximize “the homogeneity within subgroups while maximizing the heterogeneity between subgroups” (Mammadov et al., 2016, p. 177), these students might be summarized within the LPA Overachiever subgroup. Further, the Wise Operator and the Wise Learner subgroups were quite small (3.77 % and 5.05 %) and were not detected through the LPA approach. Therefore their generalizability and utility for educational practice might be questionable. It would be interesting to see if those subgroups also exist in other datasets as we can see from our data that the achievement indicators as well as the motivational-affective characteristics do differ for these groups with the cut-off score approach compared to the other high-achieving student subgroups.

Our sensitivity analysis showed that similar results would appear on an aggregated level if a narrower definition of high-achieving students (90th percentile) is used. On an individual level, the overlap decreased (47 %). This smaller overlap might be due to the subgroups that appeared only in one approach. They are, by definition, a disagreement and carry more weight in a smaller sample. Overall, both approaches showed advantages when identifying high-achieving students: The cut-

[2] When using the seven profile solution of the LPA, which would have the same number of subgroups as the cut-off score approach, the same overlap (59 %) between the two approaches on an individual level was found.

[3] When we rerun the same LPA (class-varying diagonal model) for the whole student sample, including high-achieving and not-as-high-achieving identified students ($N = 5154$), we found that for the five profile solution, the shape of two Underachiever subgroups and two Overachiever subgroups appeared. The largest and fifth subgroup was a profile with nearly the same average achievement on all three achievement indicators.

Table 6
Cross table between the cut-off score subgroups (rows) and the LPA subgroups (columns).

	Top Performer	Overachiever	Underachiever 1	Underachiever 2	Operator	Total
Top Performer	**257**	140	0	0	0	397
Overachiever	103	**235**	0	0	5	343
Underachiever	0	29	**166**	**120**	0	315
Operator	0	54	0	0	**143**	197
Learner	0	173	0	0	0	173
Wise Learner	5	69	2	3	0	79
Wise Operator	0	39	0	17	3	59
Total	365	739	168	140	151	1563

Note. The numbers presented in bold are the cases that belong to the same subgroup independent of the used methodological approach.

Table 7
Descriptive statistics of the motivational-affective characteristics for each subgroup.

	PERSEV		MATHEFF		INTMAT		INSTMOT	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Top Performer	0.35[a]	0.95	1.14	0.85	0.35	0.99	0.25	0.96
Overachiever	0.12	0.92	0.95	0.86	0.02	0.99	−0.05	0.99
Underachiever	−0.10	0.93	0.17	0.87	−0.18	1.04	−0.25	0.97
Operator	0.04	1.00	0.62	0.80	−0.37	1.04	−0.22	1.03
Learner	−0.06	0.97	0.43	0.84	−0.32	0.99	−0.37	1.02
Wise Learner	−0.06	0.86	0.39	0.91	−0.04	1.13	−0.08	0.93
Wise Operator	0.11	0.84	0.84	0.93	0.24	1.04	0.16	1.08
Top Performer	0.43	0.99	1.23	0.85	0.42	0.97	0.28	0.96
Overachiever	0.03	0.90	0.71	0.88	−0.09	1.02	−0.13	1.00
Underachiever 1	−0.18	0.97	−0.04	0.84	−0.16	1.03	−0.25	0.98
Underachiever 2	−0.01	0.88	0.38	0.79	−0.13	1.08	−0.17	0.96
Operator	0.00	1.01	0.60	0.80	−0.44	1.02	−0.30	1.06

Note. PERSEV = perseverance; MATHEFF = mathematics self-efficacy; INTMAT = mathematics interest; INSTMOT = instrumental motivation for mathematics. The upper part of the table represents the cut-off score subgroups, and the lower part the LPA subgroups.
[a] The difference from Table 8 (*M* = 0.36) is due to rounding differences in the imputed data.

Table 8
Regression analysis with each cut-off score subgroup and motivational-affective characteristics.

Subgroup	N	PERSEV			MATHEFF			INTMAT			INSTMOT		
		B	SE	t_B	B	SE	t_B	B	SE	t_B	B	SE	t_B
Constant[a]	397	0.36	0.06	6.11*	1.14	0.06	20.93*	0.35	0.06	5.74*	0.25	0.06	4.02*
Wise Operator	59	−0.24	0.14	−1.70	−0.31	0.15	−2.06*	−0.11	0.18	−0.63	−0.09	0.19	−0.47
Wise Learner	79	−0.41	0.16	−2.63*	−0.75	0.14	−5.57*	−0.39	0.17	−2.30*	−0.32	0.15	−2.12
Learner	173	−0.41	0.10	−3.96*	−0.71	0.09	−7.72*	−0.67	0.11	−6.07*	−0.62	0.11	−5.52*
Operator	197	−0.31	0.11	−2.90*	−0.53	0.10	−5.18*	−0.72	0.11	−6.60*	−0.47	0.11	−4.25*
Underachiever	315	−0.46	0.10	−4.81*	−0.98	0.08	−11.78*	−0.53	0.09	−5.71*	−0.49	0.09	−5.30*
Overachiever	343	−0.23	0.09	−2.61*	−0.20	0.08	−2.36*	−0.34	0.10	−3.50*	−0.29	0.09	−3.15*

Note. N = 1563. PERSEV = perseverance; MATHEFF = mathematics self-efficacy; INTMAT = mathematics interest; INSTMOT = instrumental motivation for mathematics. * indicates a significant result at the $\alpha = 0.05$ level after applying the Bonferroni-Holm correction.
[a] The Top Performer subgroup functions as reference group = constant.

Table 9
Regression analysis with each LPA subgroup and motivational-affective characteristics.

Subgroup	N	PERSEV			MATHEFF			INTMAT			INSTMOT		
		B	SE	t_B	B	SE	t_B	B	SE	t_B	B	SE	t_B
Constant[a]	365	0.43	0.06	6.67*	1.23	0.06	21.51*	0.42	0.06	6.71*	0.28	0.06	4.45*
Operator	151	−0.43	0.12	−3.53*	−0.63	0.11	−5.85*	−0.86	0.12	−7.45*	−0.58	0.12	−4.81*
Underachiever 1	168	−0.61	0.12	−5.25*	−1.27	0.10	−12.87*	−0.58	0.12	−4.81*	−0.54	0.12	−4.62*
Underachiever 2	140	−0.44	0.11	−3.85*	−0.85	0.10	−8.34*	−0.55	0.13	−4.40*	−0.45	0.12	−3.69*
Overachiever	739	−0.40	0.08	−5.09*	−0.52	0.07	−7.50*	−0.51	0.08	−6.52*	−0.41	0.08	−5.03*

Note. N = 1563. PERSEV = perseverance; MATHEFF = mathematics self-efficacy; INTMAT = mathematics interest; INSTMOT = instrumental motivation for mathematics. * indicates a significant result at the $\alpha = 0.05$ level after applying the Bonferroni-Holm correction.
[a] The Top Performer subgroup functions as reference group = constant.

off score approach focused on the assignment of each individual according to his or her achievement pattern. Therefore, each individual represented the overall achievement pattern of the subgroup (e.g., value above the cut-off score of the mathematical competences and value below the cut-off score of figural reasoning), and various subgroups with clear qualitative shape differences were detected. In contrast, the LPA approach created homogeneous subgroups by minimizing heterogeneity across all used achievement indicators within subgroups (Mammadov et al., 2016). Therefore, not every individual represented the overall achievement pattern of the subgroup. The advantage of the LPA approach is that students who score close to a cut-off score are not automatically split arbitrarily into two groups but can be summarized within one group or separated following quantitative level differences. Accordingly, our results showed that the Top Performers and Overachievers detected through the cut-off score approach differed clearly regarding their qualitative shape. In contrast, the Top Performers and Overachievers detected through the LPA approach differed mainly regarding their quantitative level.

Hence, if the aim would be to investigate subgroups with shape differences in achievement indicators, the cut-off score approach might be more useful. This could be relevant when qualitatively different promotion programs focus on different achievement areas, and the goal would be to assign high-achieving students to these different programs. On the other hand, the LPA approach additionally found subgroups, which differed mainly regarding their quantitative level (consistently lower vs. higher scores; cf. Morin & Marsh, 2015) across the three achievement indicators (e.g., Underachiever 1 vs. Underachiever 2 subgroup). Therefore, the LPA approach would be more helpful if the aim would be to identify high-achieving student subgroups that show quantitative level differences. This might be relevant if promotion programs focus on the same achievement area but consider different achievement levels.

In sum, our results encourage to apply multidimensional models of high achievement that include different facets of achievement (e.g., Cattell, 1943; Heller et al., 2005; Saß et al., 2017; Stoeger, 2009) and disprove the myth that high-achieving students are one homogeneous group (cf. Reis & Renzulli, 2009). This is in accordance with previous research that found remarkable achievement heterogeneity across different achievement indicators for high-achieving student subgroups (e.g., Castejón et al., 2016; Hofer & Stern, 2016; Leikin et al., 2017; Lohman et al., 2008). The positive relationships between the achievement indicators revealed from variable-centered approaches (e.g., Ehmke et al., 2017; Taub et al., 2008) did not hold true for all the different high-achieving student subgroups.

5.2. High-achieving students' motivational-affective characteristics

One possible explanation for the achievement heterogeneity are motivational-affective characteristics (e.g., Heller et al., 2005). Our results showed that high-achieving student subgroups differed regarding their motivational-affective characteristics. According to our expectations, Top Performers–which were used as a reference group because they already translated their domain-general potential into domain-specific achievement and can therefore be seen as the target state–reported the most positive motivational-affective characteristics compared to the other subgroups. The largest differences were found for mathematics self-efficacy. Here, in line with our hypothesis and previous studies (e.g., Obergriesser & Stoeger, 2015), Underachievers showed especially low scores. In turn, fostering Underachievers' self-efficacy through individualized support may help them turn their potential into achievement.

For educational practice, to help high-achieving students reach the target state of Top Performers, there are two possibilities: promoting achievement or fostering motivational-affective characteristics. As both areas are interrelated (e.g., Arens et al., 2017; Lee & Stankov, 2018), fostering one area might also have positive consequences for the other area. However, the higher malleability of motivational-affective characteristics compared to, for example, intelligence (Schonert-Reichl et al., 2017) may make this intervention path especially interesting and promising. Effective programs for students' promotion of motivational-affective characteristics already exist (cf. Mahoney et al., 2018). These interventions could be applied to high-achieving students as well. Hornstra et al. (2020) showed that promoting high-achieving students' motivation is similar to promoting average-achieving students' motivation. In line with the idea of using the development of motivational-affective characteristics as a starting point, a current meta-analysis showed that interventions for Underachievers did not necessarily improve academic achievement but consistently enhanced motivational-affective characteristics (Steenbergen-Hu et al., 2020).

Following the proposed twofold multidimensional approach–considering different achievement indicators and including motivational-affective characteristics–a promising instructional approach for supporting subgroups of (high-achieving) students in classrooms is differentiated instruction (Ziernwald et al., 2022; Deunk et al., 2018; Smale-Jacobse et al., 2019). This approach aims to proactively plan and apply different teaching methods, which take into account differences in students' abilities, learning needs, and interests (Tomlinson, 2001). With this differentiated view, the current study can help to adapt existing support programs or create new ones to help students realize their full potential. Although the knowledge about identifying and characterizing high-achieving students is an important and necessary prerequisite to adapt teaching appropriately (e.g., Castejón et al., 2016), it is not sufficient. Therefore, future studies should investigate from which teaching strategies the different high-achieving student subgroups benefit the most.

5.3. Limitations and future research

As the present study focused exclusively on mathematics for the domain-specific achievement indicators, it remains an open question whether the results can be applied to other domains. For the general relationship between achievement in different domains, Holenstein et al. (2021) showed in a longitudinal study with secondary students that mathematical competence was related to academic achievement in various school domains, such as reading comprehension and scientific literacy, when tested four years later. Another study focusing on science, which defined high-achieving students as students scoring in the top 20 % in PISA 2012, found that 56 % of these students showed high mathematics, science, and reading achievement, and further 32 % showed high achievement in science and one other domain (Schmidtner, 2017). Accordingly, we would assume that quite a high overlap between the different achievement domains exists, and high-achieving students, to a great extent, show high achievement in different domains and not only in one domain. For the identification of different subgroups, we would assume that similar subgroups would have been detected regardless of the achievement domain, as the focus of the subgroups is less on the specific content of mathematical competence but more on domain-general vs. application-oriented vs. curriculum-oriented competence. However, future studies may investigate to what extent these assumptions can be empirically supported by including achievement indicators from different domains. In addition, for analyzing the relationship between high-achieving students' potential and their current achievement, additional motivational-affective characteristics such as self-concept and achievement goals could be included. Hints that there might be differences between high-achieving student subgroups in self-concept and achievement goals (e.g., master goals) have been shown in previous research comparing high-achieving versus non-high-achieving gifted students (e.g., Lüftenegger et al., 2015).

Further, the current study is based on cross-sectional data. Therefore, no transitions between subgroups could be investigated. Empirical studies showed that mathematical competence, as well as intelligence, are quite stable over a longer period of time (Holenstein et al. (2021); Larsen et al., 2008; Yu et al., 2018), which leads to the assumption that the detected subgroups might also be stable over a longer period of time. However, especially through interventions, motivational-affective characteristics–that are (reciprocally) connected with achievement (e.g., Arens et al., 2017; Lee & Stankov, 2018)–could change (e.g., Steenbergen-Hu et al., 2020). Therefore, transitions between subgroups later on are also conceivable. Future studies may examine the identification and development of high-achieving student subgroups by investigating achievement indicators and motivational-affective characteristics across time.

What should be taken into account when interpreting the achievement level reached in the current study is that testing was under low-stakes conditions. An advantage of low-stakes testing might be that test anxiety is lower compared to high-stakes testing, and therefore, achievement can be measured more appropriately (cf. Finn, 2015). Further, in our study, three-quarters of the high-achieving students indicated that they were trying the PISA test just as hard or just a little less hard compared to a personally highly important situation. However, it may be that students would have shown more effort in high-stakes testing, which in turn may have been reflected in somewhat higher achievement and more students classified as high-achieving and especially as Top Performers. Further, a decline in test motivation may have also influenced the BEFKI achievement, measured at the end of the second test day. For example, Lindner et al. (2019) showed that increasing testing time raised the probability of guessing behavior. This possible decline in motivation might have resulted in an overestimation of Overachievers. However, the authors also showed that high-achieving students were less likely to start guessing at an early time than students with lower achievement.

6. Conclusion

The present study's added value is (a) a direct comparison of two methodological approaches, namely the cut-off score and the LPA approach, (b) the use of multiple achievement indicators, and (c) the examination of motivational-affective characteristics. Both approaches revealed the four largest high-achieving student subgroups defined through the cut-off score approach, namely Top Performers, Overachievers, Underachievers, and Operators. The heterogeneity of high-achieving students could be unraveled more by the comprehensive consideration of domain-general and domain-specific achievement indicators as well as motivational-affective characteristics. In order to support high-achieving students in transferring their potential into achievement, our results suggest that both cognitive competences and motivational-affective characteristics should be taken into account. Hence, the study emphasizes the importance of a multidimensional understanding of high-achieving students. This may help to raise the awareness of researchers, teacher educators, and teachers that learning is a highly individual process and that, thereby, for teaching and promotion, it is not enough to distinguish between high-, average-, and low-achieving students. Only when there is an awareness of this large achievement and motivational-affective heterogeneity–also within high-, average- or low-achieving student subgroups–can students be educated in the best possible way according to their individual needs.

Declaration of competing interest

None.

Appendix A. Achievement indicators' main characteristics

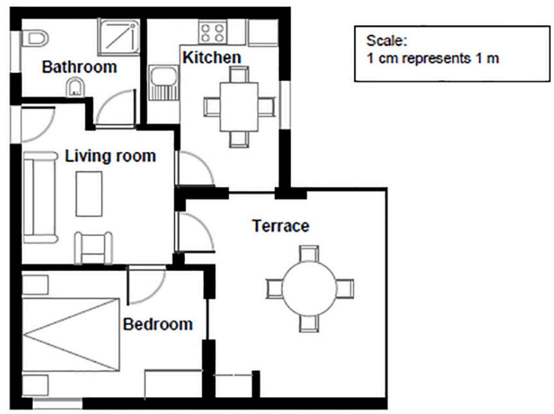
Achievement indicator	Main characteristics
Fluid intelligence (Baumert et al., 2009; Cattell, 1943; Schulze et al., 2005)	• Domain-general achievement • Focus on decontextualized knowledge without content-specific knowledge • Measured through the deductive reasoning subscale (BEFKI)
Curriculum-oriented mathematical competence (Hartig & Frey, 2012; KMK, 2016b; Pant et al., 2013a; Roppelt, Blum, & Pöhlmann, 2013)	• Domain-specific achievement • Focus on curricular competences and contains next to application-oriented tasks also tasks without real-world context (purely mathematical tasks) • Measured through the mathematics global scale in NAS
Application-oriented mathematical competence (de Lange, 2006; OECD, 2013a)	• Domain-specific achievement • Focus on the application of acquired mathematical knowledge in everyday contexts (i.e., personal, occupational, societal, and scientific) • Measured through the mathematics global scale in PISA

Appendix B. Released sample items for mathematical competences in PISA and NAS

PISA (adapted from OECD, 2013b)

APARTMENT PURCHASE

This is the plan of the apartment that George's parents want to purchase from a real estate agency.



Question 1: APARTMENT PURCHASE PM00FQ01 – 0 1 9

To estimate the total floor area of the apartment (including the terrace and the walls), you can measure the size of each room, calculate the area of each one and add all the areas together.

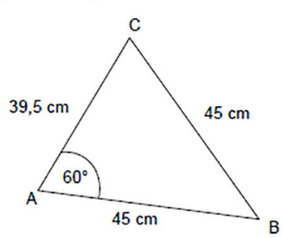
However, there is a more efficient method to estimate the total floor area where you only need to measure 4 lengths. Mark on the plan above the **four** lengths that are needed to estimate the total floor area of the apartment.

NAS (translated from IQB, 2018)

Released Sample Mathematics Items IQB Trends in Student Achievement

Proficiency Area 2: Measurements

Task: " The impossible triangle"



Give a reason why there cannot be a triangle with these dimensions.

Appendix C. Correlations of the achievement indicators for each subgroup

Subgroup	Correlation PISA & NAS	Correlation PISA & BEFKI	Correlation NAS & BEFKI
Wise operator	−0.09	−0.09	0.01
Wise learner	0.29*	0.25*	0.23*
Learner	0.13	0.14	0.04
Operator	0.02	−0.11	0.18*
Underachiever	0.58*	0.03	0.10
Overachiever	0.46*	0.05	0.17*
Top performer	0.49*	0.08	0.22*
Operator	−0.12	−0.15	−0.08
Underachiever 1	0.28*	0.12	0.19*
Underachiever 2	−0.06	0.17*	−0.06
top performer	0.15*	−0.07	0.17*
Overachiever	0.04	−0.02	−0.11*

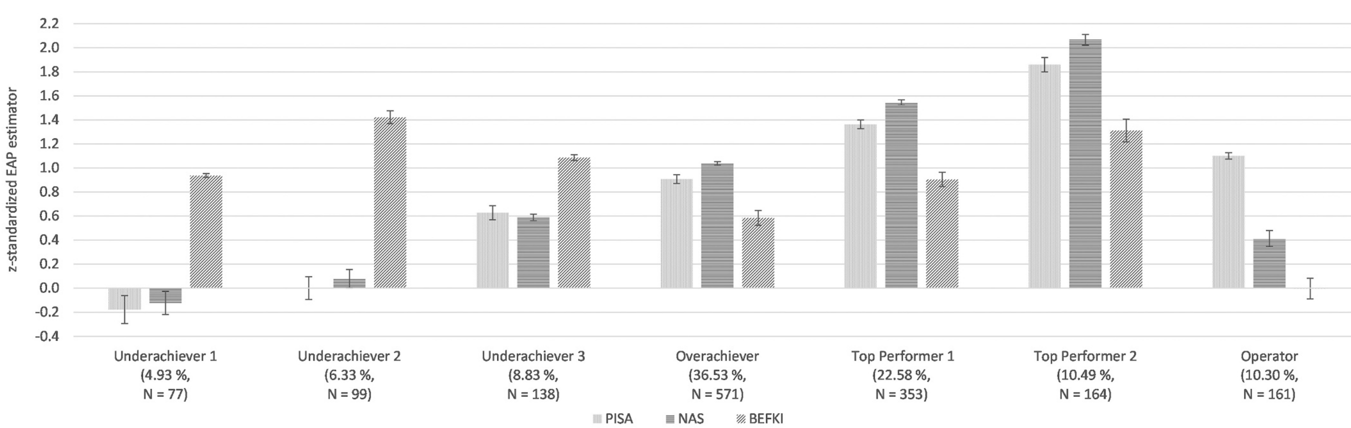
Note. PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning. The upper part of the table represents the cut-off score subgroups and the lower part the LPA subgroups.

* $p < .05$.

Appendix D. Percentage of girls and students attending the highest school track (gymnasium) for each subgroup

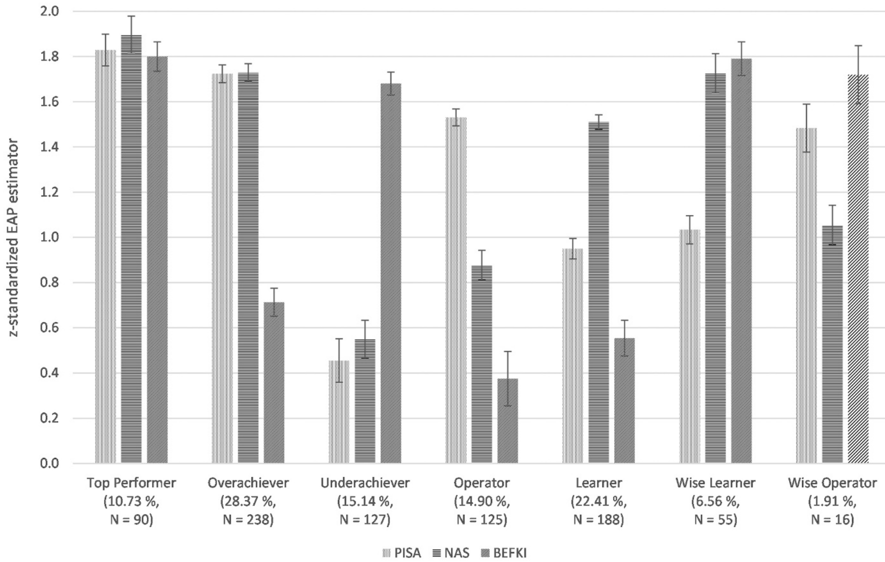
Cut-off score subgroup	Percentage girls	Percentage highest school track	LPA subgroup	Percentage girls	Percentage highest school track
Underachiever	58.4 %	47.3 %	Underachiever 1	61.3 %	32.7 %
Wise Learner	59.5 %	73.4 %	Underachiever 2	57.9 %	63.6 %
Wise Operator	50.8 %	74.6 %	Operator	31.1 %	74.2 %
Operator	33.5 %	76.1 %	Overachiever	47.2 %	84.4 %
Learner	55.5 %	83.8 %	Top Performer	36.2 %	95.9 %
Top Performer	41.1 %	91.7 %			
Overachiever	36.7 %	93.3 %			

Appendix E. Means (±2 SE) of the achievement indicators for each of the 7 LPA subgroups

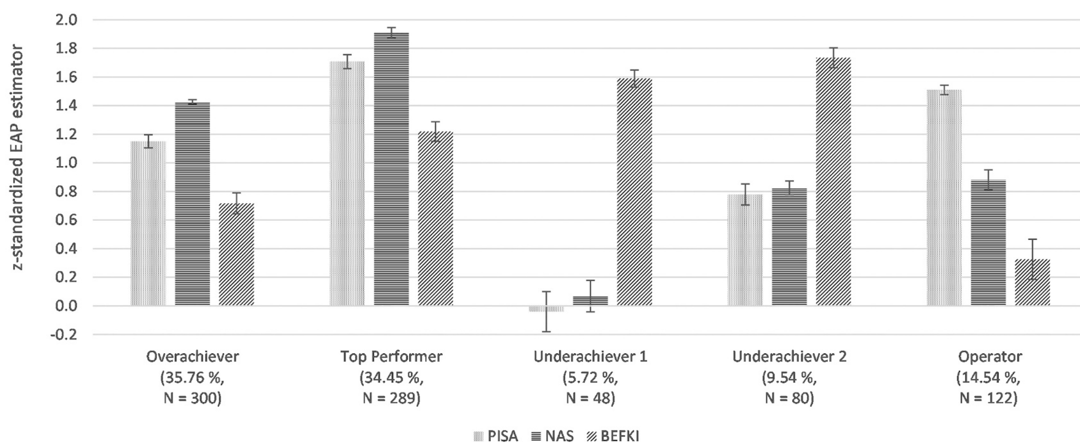


Note. PISA = application-oriented mathematics; NAS = curriculum-oriented mathematics; BEFKI = figural reasoning.

Appendix F. Means (±2 SE) of the achievement indicators for each cut-off score subgroup with the 90th percentile sample



Appendix G. Means (±2 SE) of the achievement indicators for each of the 5 LPA subgroups with the 90th percentile sample



References

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

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621–634. https://doi.org/10.1037/edu0000163

Ashcraft, M. H., & Rudig, N. O. (2012). Higher cognition is altered by noncognitive factors: How affect enhances and disrupts mathematics performance in adolescence and young adulthood. In V. F. Reyna, S. B. Chapman, M. R. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 243–263). American Psychological Association. https://doi.org/10.1037/13493-009.

Bacher, J., Pöge, A., & Wenzig, K. (2010). *Clusteranalyse: Anwendungsorientierte Einführung in Klassifikationsverfahren [Cluster analysis: Application-oriented introduction to classification methods]*. Oldenbourg.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in Changing Societies* (pp. 1–45). Cambridge University Press.

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

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology, 39*(2), 174–187. https://doi.org/10.1093/jpepsy/jst084

Cappelli, P. H. (2015). Skill gaps, skill shortages, and skill mismatches. *ILR Review, 68*(2), 251–290. https://doi.org/10.1177/0019793914564961

Carman, C. A. (2013). Comparing apples and oranges: Fifteen years of definitions of giftedness in research. *Journal of Advanced Academics, 24*(1), 52–70. https://doi.org/10.1177/1932202X12472602

Castejón, J. L., Gilar, R., Miñano, P., & González, M. (2016). Latent class cluster analysis in exploring different profiles of gifted and talented students. *Learning and Individual Differences, 50*, 166–174. https://doi.org/10.1016/j.lindif.2016.08.003

Cattell, R. B. (1943). The measurment of adult intelligence. *Psychological Bulletin, 40*(3), 153–193.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *The Journal of General Psychology, 132*(1), 5–40. https://doi.org/10.3200/GENP.132.1.5-40

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.

Cross, T. L., & Cross, J. R. (2017). Maximizing potential: A school-based conception of psychosocial development. *High Ability Studies, 28*(1), 43–58. https://doi.org/10.1080/13598139.2017.1292896

Cutumisu, M., & Bulut, O. (2017). Problem-solving attitudes and gender as predictors of academic achievement in mathematics and science for Canadian and Finnish students in the PISA 2012 assessment. *Journal of Educational Multimedia and Hypermedia, 26*(4), 325–342.

de Jonge, H., Jak, S., & Kan, K.-J. (2020). Dealing with artificially dichotomized variables in meta-analytic structural equation modeling. *Zeitschrift für Psychologie, 228*(1), 25–35. https://doi.org/10.1027/2151-2604/a000395

de Lange, J. (2006). Mathematical literacy for living from OECD-PISA perspective. *Tsukuba Journal of Educational Study in Mathematics, 25*, 13–35.

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

Ehmke, T., Köller, O., & Stanat, P. (2017). Äquivalenz der Erfassung mathematischer Kompetenzen in PISA 2012 und im IQB-Ländervergleich 2012 [Equivalence in measuring mathematical competence in PISA 2012 and the National Assessment 2012]. *Zeitschrift für Erziehungswissenschaft, 20*(S2), 37–59. https://doi.org/10.1007/s11618-017-0751-5

Ehmke, T., van den Ham, A.-K., Sälzer, C., Heine, J., & Prenzel, M. (2020). Measuring mathematics competence in international and national large scale assessments: Linking PISA and the national educational panel study in Germany. *Studies in Educational Evaluation, 65*, Article 100847. https://doi.org/10.1016/j.stueduc.2020.100847

Enders, C. K. (2010). *Applied missing data analysis. Methodology in the social sciences*. Guilford Press.

Finn, B. (2015). Measuring motivation in low-stakes assessments. *ETS Research Report Series, 15*(19), 1–17. https://doi.org/10.1002/ets2.12067

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

Fung, F., Tan, C. Y., & Chen, G. (2018). Student engagement and mathematics achievement: Unraveling main and interactive effects. *Psychology in the Schools, 55*(7), 815–831. https://doi.org/10.1002/pits.22139

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies, 15*(2), 119–147. https://doi.org/10.1080/1359813042000314682

Geiser, C. (2013). *Data analysis with Mplus*. The Guilford Press.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science : The Official Journal of the Society for Prevention Research, 8*(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Grimm, K. J., Mazza, G. L., & Davoudzadeh, P. (2017). Model selection in finite mixture models: A k-fold cross-validation approach. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(2), 246–256. https://doi.org/10.1080/10705511.2016.1250638

Gubbels, J., Segers, E., & Verhoeven, L. (2018). How children's intellectual profiles relate to their cognitive, socio-emotional, and academic functioning. *High Ability Studies, 29*(2), 149–168. https://doi.org/10.1080/13598139.2018.1507902

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling, 25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hartig, J., & Frey, A. (2012). Validität des Tests zur Überprüfung des Erreichens der Bildungsstandards in Mathematik. Zusammenhänge mit den bei PISA gemessenen Kompetenzen und Varianz zwischen Schulen und Schulformen [Validity of a standard-based test for mathematical competencies. Relations with the competencies assessed in PISA and variance between schools and school tracks]. *Diagnostica, 58*(1), 3–14. https://doi.org/10.1026/0012-1924/a000064

Hecht, M., Roppelt, A., & Siegle, T. (2013). Testdesign und Auswertung des Ländervergleichs [Test design and analysis of the National Assessment Study]. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle, & C. Pöhlmann (Eds.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (pp. 391–402). Waxmann.

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. https://doi.org/10.1017/CBO9780511610455.010.

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences, 66*, 4–15. https://doi.org/10.1016/j.lindif.2017.11.001

Hidi, S., & Renninger, A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hofer, S. I., & Stern, E. (2016). Underachievement in physics: When intelligent girls fail. *Learning and Individual Differences, 51*, 119–131. https://doi.org/10.1016/j.lindif.2016.08.006

Holenstein, M., Bruckmaier, G., & Grob, A. (2021). Transfer effects of mathematical literacy: An integrative longitudinal study. *European Journal of Psychology of Education, 36*, 799–825. https://doi.org/10.1007/s10212-020-00491-4

Hong, E., & Aqui, Y. (2004). Cognitive and motivational characteristics of adolescents gifted in mathematics: Comparisons among students with different types of giftedness. *Gifted Child Quarterly, 48*(3), 191–201. https://doi.org/10.1177/001698620404800304

Hong, W., Bernacki, M. L., & Perera, H. N. (2020). A latent profile analysis of undergraduates' achievement motivations and metacognitive behaviors, and their relations to achievement in science. *Journal of Educational Psychology, 112*(7), 1409–1430. https://doi.org/10.1037/edu0000445

Hornstra, L., Bakx, A., Mathijssen, S., & Denissen, J. J. (2020). Motivating gifted and non-gifted students in regular primary schools: A self-determination perspective. *Learning and Individual Differences, 80*. https://doi.org/10.1016/j.lindif.2020.101871. preprint.

Howard, M. C., & Hoffman, M. E. (2018). Variable-centered, person-centered, and person-specific approaches. *Organizational Research Methods, 21*(4), 846–876. https://doi.org/10.1177/1094428117744021

Institute of Medicine. (2007). Rising above the gathering storm: Energizing and employing America for a brighter economic future. *National Academies Press*. https://doi.org/10.17226/11463

IQB. (2018). Beispielaufgaben Mathematik – Das unmögliche Dreieck [Released Sample Mathematics Items IQB Trends in Student Achievement – The impossible triangle]. https://www.iqb.hu-berlin.de/bt/bt/BT2018/Beispielaufgaben/DasunmglicheDrei.pdf.

Klieme, E., Neubrand, M., & Lüdtke, O. (2001). Mathematische grundbildung: Testkonzeption und ergebnisse [Mathematical literacy: Test conception and results]. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, … M. Weiß (Eds.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 139–190). Leske + Budrich.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education, 33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Kuger, S., & Klieme, E. (2016). Dimensions of context assessment. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning* (pp. 3–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-45357-6_1.

Kultusministerkonferenz. (2016a). Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler [Joint initiative of the Federal Ministry of Education and Research and the Standing Conference of Ministers of Education and Cultural Affairs to promote high-achieving and potentially high-achieving students]. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2016/2016-11-28_Gem.Initiative_Leistungsstarke-Beschluss.pdf.

Kultusministerkonferenz. (2016b). Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring [Overall strategy of the Standing Conference of Ministers of Education and Cultural Affairs on educational monitoring]. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf.

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle-age. *Intelligence, 36*(1), 29–34. https://doi.org/10.1016/j.intell.2007.01.001

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin Company.

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

Lehner, M. C., Heine, J.-H., Sälzer, C., Reiss, K., Haag, N., & Heinze, A. (2017). Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe [Changes in mathematical competence from grade 9 to grade 10]. *Zeitschrift für Erziehungswissenschaft, 20*(S2), 7–36. https://doi.org/10.1007/s11618-017-0746-2

Leikin, R., Leikin, M., Paz-Baruch, N., Waisman, I., & Lev, M. (2017). On the four types of characteristics of super mathematically gifted students. *High Ability Studies, 28*(1), 107–125. https://doi.org/10.1080/13598139.2017.1305330

Leikin, R., Paz-Baruch, N., & Leikin, M. (2014). Cognitive characteristics of students with superior performance in mathematics. *Journal of Individual Differences, 35*(3), 119–129. https://doi.org/10.1027/1614-0001/a000140

Lens, W., Paixão, M. P., & Herrera, D. (2009). Instrumental motivation is extrinsic motivation: So what??? *Psychologica, 50*, 21–40. https://doi.org/10.14195/1647-8606_50_2

Lenski, A. E., Hecht, M., Penk, C., Milles, F., Mezger, M., Heitmann, P., Stanat, P., & Pant, H. A. (2016). *IQB-Ländervergleich 2012. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente [The IQB National Assessment Study 2012. Scale manual for the documention of survey instruments]*. Humboldt-Universität zu Berlin: Institut zur Qualitätsentwicklung im Bildungswesen. https://doi.org/10.20386/HUB-42547

Lindner, M. A., Lüdtke, O., & Nagy, G. (2019). The onset of rapid-guessing behavior over the course of testing time: A matter of motivation and cognitive resources. *Frontiers in Psychology, 10*, 1533. https://doi.org/10.3389/fpsyg.2019.01533

Litster, K., & Roberts, J. (2011). The self-concepts and perceived competencies of gifted and non-gifted students: A meta-analysis. *Journal of Research in Special Educational Needs, 11*(2), 130–140. https://doi.org/10.1111/j.1471-3802.2010.01166.x

Lohman, D. F., Gambrell, J., & Lakin, J. (2008). The commonality of extreme discrepancies in the ability profiles of academically gifted students. *Psychology Science Quarterly, 50*(2), 269–282.

Lüftenegger, M., Kollmayer, M., Bergsmann, E., Jöstl, G., Spiel, C., & Schober, B. (2015). Mathematically gifted students and high achievement: The role of motivation and

classroom structure. *High Ability Studies, 26*(2), 227–243. https://doi.org/10.1080/13598139.2015.1095075
Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan, 100*(4), 18–23. https://doi.org/10.1177/0031721718815668
Mammadov, S., Ward, T. J., Cross, J. R., & Cross, T. L. (2016). Use of latent profile analysis in studies of gifted students. *Roeper Review, 38*(3), 175–184. https://doi.org/10.1080/02783193.2016.1183739
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
Mielicki, M. K., Fitzsimmons, C. J., Woodbury, L. H., Marshal, H., Zhang, D., Rivera, F. D., & Thompson, C. A. (2021). Effects of figural and numerical presentation formats on growing pattern performance. *Journal of Numerical Cognition, 7*(2), 125–155. https://doi.org/10.5964/jnc.6945
Morin, A. J. S., & Marsh, H. W. (2015). Disentangling shape from level effects in person-centered analyses: An illustration based on university teachers' multidimensional profiles of effectiveness. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(1), 39–59. https://doi.org/10.1080/10705511.2014.919825
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.
Neidorf, T. S., Binkley, M., Gattis, K., & Nohara, D. (2006). Comparing mathematics content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 assessments: Technical report (NCES 2006-029). Institute of Education Sciences. https://nces.ed.gov/pubs2006/2006029.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
Neuendorf, C., Kuhl, P., & Jansen, M. (2017). Leistungsstarke Schülerinnen und Schüler in Deutschland [high-achieving students in Germany]. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich, & N. Haag (Eds.), *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (pp. 317–334). Waxmann.
Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569. https://doi.org/10.1080/10705510701575396
Obergriesser, S., & Stoeger, H. (2015). The role of emotions, motivation, and learning behavior in underachievement and results of an intervention. *High Ability Studies, 26*(1), 167–190. https://doi.org/10.1080/13598139.2015.1043003
Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson, & M. Kaptein (Eds.), *Human-computer interaction series. Modern statistical methods for HCI* (pp. 275–287). Springer International Publishing.
OECD. (2009). *PISA data analysis manual* (2nd ed.). OECD.
OECD. (2013a). *PISA 2012 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/9789264190511-en
OECD. (2013b). PISA 2012 released mathematics items. https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf.
OECD. (2013c). *PISA 2012 results: Ready to learn: Students` engagment, drive and self-beliefs. III*. OECD Publishing. https://doi.org/10.1787/9789264201170-en
OECD. (2014a). *PISA 2012 results: What students know and can do - Student performance in mathematics, reading and science* ((Revised Edition)) ((Revised Edition), Volume I). OECD Publishing.
OECD. (2014b). *PISA 2012 technical report*. OECD.
OECD. (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/b25efab8-en
OECD. (2020). PISA 2024: Strategic vision and direction for science. https://www.oecd.org/pisa/publications/PISA-2024-Science-Strategic-Vision-Proposal.pdf.
Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (2013a). The IQB National Assessment Study 2012: Competencies in mathematics and the sciences at the end of secondary level I. https://www.iqb.hu-berlin.de/bt/lv2012/Bericht.
Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (Eds.). (2013b). *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I [The IQB National Assessment Study 2012: Competencies in mathematics and the sciences at the end of secondary level I]. Waxmann*.
Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32*(1), 8–47. https://doi.org/10.1016/j.cedpsych.2006.10.003
Pitsia, V., Biggart, A., & Karakolidis, A. (2017). The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement: A multilevel analysis of the programme for international student assessment data. *Learning and Individual Differences, 55*, 163–173. https://doi.org/10.1016/j.lindif.2017.03.014
Preckel, F., Golle, J., Grabner, R., Jarvin, L., Kozbelt, A., Müllensiefen, D., Olszewski-Kubilius, P., Schneider, W., Subotnik, R., Vock, M., & Worrell, F. C. (2020). Talent development in achievement domains: A psychological framework for within- and cross-domain research. *Perspectives on Psychological Science, 15*(3), 691–722. https://doi.org/10.1177/1745691619895030
Reis, S. M., & Renzulli, J. S. (2009). Myth 1: The gifted and talented constitute one single homogeneous group and giftedness is a way of being that stays in the person over time and experiences. *Gifted Child Quarterly, 53*(4), 233–235. https://doi.org/10.1177/0016986209346824
Renzulli, J. S. (1984). The Triad/ revolving door system: A research-based approach to identification and programming for the gifted and talented. *Gifted Child Quarterly, 28*(4), 163–171. https://doi.org/10.1177/001698628402800405
Ritchotte, J. A., Matthews, M. S., & Flowers, C. P. (2014). The validity of the achievement-orientation model for gifted middle school students. *Gifted Child Quarterly, 58*(3), 183–198. https://doi.org/10.1177/0016986214534890
Roppelt, A., Blum, W., & Pöhlmann, C. (2013). Beschreibung der untersuchten mathematischen Kompetenzen [Description of the examined mathematical competencies]. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle, & C. Pöhlmann (Eds.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (pp. 23–37). Waxmann.
Roppelt, A., Penk, C., Pöhlmann, C., & Pietsch, E. (2013). Der Ländervergleich im Fach Mathematik. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle, & C. Pöhlmann (Eds.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (pp. 123–140). Waxmann.
Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion [Textbook test theory – test construction]* (Vol. 2). Psychologie Lehrbuch: Hans Huber.
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020
Saß, S., Kampa, N., & Köller, O. (2017). The interplay of g and mathematical abilities in large-scale assessments across grades. *Intelligence, 63*, 33–44. https://doi.org/10.1016/j.intell.2017.05.001
Schiepe-Tiska, A. (2019). School tracks as differential learning environments moderate the relationship between teaching quality and multidimensional learning goals in mathematics. *FrontiersEducation, 4*(4). https://doi.org/10.3389/feduc.2019.00004
Schmidtner, S. S. (2017). *Leistungsstarke Jugendliche in Naturwissenschaften: Vertiefende Analysen zu PISA 2006 und PISA 2012 [High performing students in science: In-depth analyses of PISA 2006 and PISA 2012]*. München: Technische Universität München [Dissertation].
Schonert-Reichl, K. A., Kitil, M. J., & Hanson-Peterson, J. (2017). *To reach the students, teach the teachers: A national scan of teacher preparation and social and emotional learning*. Vancouver, B.C.
Schulze, D., Beauducel, A., & Brocke, B. (2005). Semantically meaningful and abstract figural reasoning in the context of fluid and crystallized intelligence. *Intelligence, 33*(2), 143–159. https://doi.org/10.1016/j.intell.2004.07.011
Siegle, T., Schroeders, U., & Roppelt, A. (2013). Anlage und Durchführung des Ländervergleichs [Setup and implementation of the National Assessment Study]. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle, & C. Pöhlmann (Eds.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (pp. 101–121). Waxmann.
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
Steenbergen-Hu, S., Olszewski-Kubilius, P., & Calvert, E. (2020). The effectiveness of current interventions to reverse the underachievement of gifted students: Findings of a meta-analysis and systematic review. *Gifted Child Quarterly, 64*(2), 132–165. https://doi.org/10.1177/0016986220908601
Stoeger, H. (2009). The history of giftedness research. In L. V. Shavinina (Ed.), *International handbook on giftedness: Part one* (pp. 17–38). Netherlands: Springer. https://doi.org/10.1007/978-1-4020-6162-2_2.
Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest : A Journal of the American Psychological Society, 12*(1), 3–54.
Taskinen, P. H., Schütte, K., & Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: The role of school factors, individual interest, and science self-concept. *Educational Research and Evaluation, 19*(8), 717–733. https://doi.org/10.1080/13803611.2013.853620
Taub, G. E., Keith, T. Z., Floyd, R. G., & Mcgrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly, 23*(2), 187–198. https://doi.org/10.1037/1045-3830.23.2.187
Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Association for Supervision & Curriculum Development.
Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8-10) [Berliner Test for the acquisition of fluid and crystallized intelligence in Grades 8 to 10]. Hogrefe*.
Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence, 69*, 8–15. https://doi.org/10.1016/j.intell.2018.03.011
Ziernwald, L., Hillmayr, D., & Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms—A systematic review. *Journal of Advanced Academics, 33*(4), 540–573. https://doi.org/10.1177/1932202X221112931

Appendix B - Paper B

Ziernwald, L., Hillmayr, D., & Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms–A systematic review. *Journal of Advanced Academics*, *33*(4), 540–573. https://doi.org/10.1177/1932202X221112931

Article

Journal of Advanced Academics
2022, Vol. 33(4) 540–573

DOI: 10.1177/1932202X221112931
journals.sagepub.com/home/joaa
SAGE

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

Lisa Ziernwald[1], Delia Hillmayr[1], and Doris Holzberger[1]

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

[1]Centre 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 subgroups (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 (Agaliotis & 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.

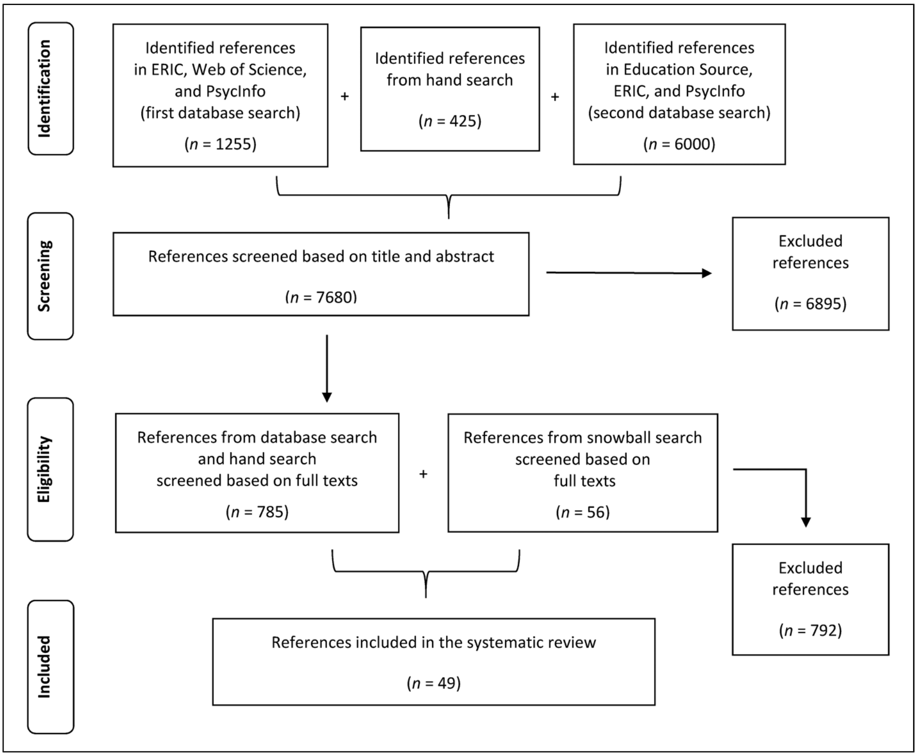


Figure 1. Flow diagram describing the study search and selection process.

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 studies[1] (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, $n = 25$)		
	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, $n = 15$[a])		
	Mathematics	6
	Science	4
	Reading	4
	Motivational-affective outcomes	6
	Other	3
Teacher training regarding DI (only studies investigating DI's impact, $n = 15$)		
	Yes	10
	No	5

[a]As 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_{ppc2}	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
				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
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 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_{ppc2}	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	1.127	0.058
				Science content knowledge year 1	3.155	0.121
				Science concept knowledge year 1	2.089	0.077
				Science process skills year 2	0.426	0.029
				Science content knowledge year 2	7.179	0.353
				Science concept knowledge year 2	0.856	0.044
Saleh et al., 2005	Grouping: Homogeneous grouping	Heterogeneous grouping	Standardized achievement Test: upper 25% on Science Elementary Achievement Test (SEAT)	Science achievement: biology	0.725	0.121
				Motivational beliefs 1: academic benefits	0.483	0.116
				Motivational beliefs 2: social benefits	−0.149	0.112
				Motivational beliefs 3: attitude benefits	−0.031	0.112

(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]
Shaunessy-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	0.194	0.141
				Attitudes toward reading	−0.068	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.
[a]Sampling variances were calculated according to Viechtbauer (2021). Because only one study gave the correlation between pre- and posttest (Shaunessy-Dedrick et al., 2015), we used $r = .60$ as a conservative estimate for the studies that did not report the correlation (cf. van Alten 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 (RQ 3)

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

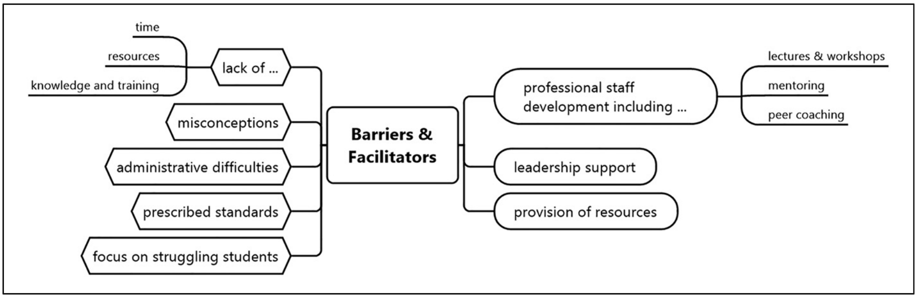


Figure 2. Barriers and facilitators for implementing differentiated instruction for high-achieving students in mixed-ability classrooms. Note. The hexagons on the left side represent barriers, and the ovals on the right side represent facilitators for implementing differentiated instruction.

(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., Gunnulfsen, 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/10.1177/1094428 106291059

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' engagment, 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 Jugendlich [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/10.1177/001698 6210383155

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

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/10.1177/01623532 0502900104

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 innovatitve 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.