Elementary School Teachers' Beliefs and Judgments About Students' Giftedness

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

Dipl.-Psych. Sandra Rothenbusch

aus Erfurt

Tübingen

2016

Tag der mündlichen Prüfung:	20.12.2016
Dekan:	Professor Dr. rer. soc. Josef Schmid
1. Gutachter:	Associate Professor Dr. phil. Ingo Zettler
2. Gutachter:	Professorin Dr. phil. Thamar Voss
3. Gutachter:	Professor Dr. phil. Ulrich Trautwein

ACKNOWLEDGMENTS

This dissertation arose within the context of the project "Formative Assessment of the Hector Children's Academy Program" and was sustained by a scholarship through the Landesgraduiertenförderung (LGF) at the Eberhard Karls Universität Tübingen and by grants from the Hector Foundation II. I sincerely thank them all for the possibility to realize this work.

I would like to express my deep gratitude to both my supervisors, Assoc. Prof. Dr. Ingo Zettler and Prof. Dr. Thamar Voss, for their support, guidance, and feedback throughout all stages of the dissertation. Thank you!

Special thanks go out to my former colleagues at the Hector Research Institute of Education Sciences and Psychology at the Eberhard Karls Universität Tübingen—in particular the "Hector research team"—for the inspiring research environment and their support. Furthermore, I would like to cordially thank all my colleagues at the Institute for Educational Psychology at the Leibniz Universität Hannover for supporting me in my long last stage of this dissertation.

Finally, I owe many thanks to my husband Stefan Rothenbusch who has been constantly there for me, who supported me emotionally, and who has always been on call when I needed help to express myself in English. I also would like to say a big thank you to my family for believing in me.

ABSTRACT

Teachers play a significant role in the academic development of gifted elementary school students (e.g., Brighton, Moon, Jarvis, & Hockett, 2007; Rost & Schilling, 2006). Specifically, their judgments about which students from their classrooms are gifted—and therefore those who have the potential for or show excellent achievement (see Sternberg & Davidson, 2005)—are important, for example, when they are involved in the identification process for gifted education programs. The present dissertation deals with elementary school teachers' beliefs and judgments about giftedness: what they believe giftedness is, whom they see as gifted, and how they judge facets of giftedness among students who they see as gifted. By means of three empirical studies, beliefs about giftedness were related to a scientific conception of giftedness and judgments were connected to teacher variables, with student characteristics on the individual and class levels, and with parent judgments.

Study 1 explored the question of which beliefs teachers hold about giftedness. Based on a modern conception of giftedness, eight dimensions for beliefs about the content and development of giftedness were derived. Four groups were compared: student teachers, elementary school teachers, and teachers of a gifted education program of which some were also school teachers. Additionally, respondents' beliefs were set into relation with the number of years in general classrooms and in the gifted education program as well as with beliefs about the malleability of intelligence. The theoretically derived dimensions could be empirically supported. Notwithstanding the huge diversity of beliefs, teachers tended to agree with the conception of giftedness used. Partial measurement invariance was achieved between the groups. The beliefs of student teachers differed the most from those of the other teachers, whereas the latter were similar to each other. No or few differences in beliefs were found for years working in a general classroom or gifted education program. However, beliefs about giftedness were related to beliefs about intelligence.

Study 2 explored the question of which elementary school students got nominated by teachers for a gifted education program. It was hypothesized that the probability of getting nominated was positively related to students' individual intelligence, but also that a negative reference group effect would occur. Hence, the probability of getting nominated should also be negatively associated with the average level of intelligence in

a class. Furthermore, the effects of experience in the area of giftedness, beliefs on whether giftedness is holistic or domain-specific, and beliefs on whether intelligence is malleable or fixed on the reference group effect were investigated. As expected, students' individual intelligence was positively associated with the probability of getting nominated. Also, support was found for the negative reference group effect: Students had higher probabilities of getting nominated when they were in classes with lower average levels of intelligence than students with similar intelligence scores who were in classes with higher average levels of intelligence. The negative reference group effect was stronger for teachers who saw giftedness as holistic instead of domain-specific. Teacher beliefs about intelligence had no effect, but different kinds of experience in the area of giftedness were differently connected to the reference group effect.

With *Study 3*, teacher judgments of nominated students were compared with parent judgments and the effect of congruence between both judgments on students' German and math grades was investigated. Judgments about verbal and mathematical abilities, deductive reasoning, creative thinking, and engagement were measured. Teachers and parents were compared concerning their judgmental accuracy levels and whether their judgments were affected by halo effects. Furthermore, the congruence between teacher and parent judgments was determined. The accuracy levels of teacher and parent judgments did not differ from each other. Both judgments were affected by halo effects, but teachers were more affected than parents. The congruence between teacher and parent judgments was low to mediocre overall. High teacher and parent judgments were related to better German grades in an additive matter. Higher parent judgments of mathematical abilities and engagement reduced the associations between teacher judgments and math grades.

In the general discussion, the results of the three studies are summarized and critically discussed. Furthermore, implications for research and educational praxis are deduced.

ZUSAMMENFASSUNG

Lehrkräfte nehmen eine wichtige Rolle in der akademischen Entwicklung von hochbegabten Grundschulkindern ein (z.B. Brighton, Moon, Jarvis & Hockett, 2007; Rost & Schilling, 2006). Insbesondere ihre Einschätzung, welche Kinder in ihrem Klassenraum hochbegabt sind – also das Potential zu außerordentlicher Leistung haben oder ungewöhnlich leistungsstark sind (siehe Sternberg & Davidson, 2005) – ist von Bedeutung, beispielsweise wenn sie in den Identifikationsprozess von Kindern für Hochbegabtenförderprogramme eingebunden sind. Die vorliegende Dissertation beschäftigt sich mit Hochbegabungsüberzeugungen und -einschätzungen von Grundschullehrkräften: was sie unter Hochbegabung verstehen, wen sie als hochbegabt ansehen und wie sie hochbegabungsrelevante Merkmale dieser Schülerinnen und Schüler einschätzen. Mit Hilfe von drei empirischen Studien wurden die Überzeugungen mit einer wissenschaftlichen Hochbegabungskonzeption und die Einschätzungen mit Lehrermerkmalen, Schülermerkmalen auf der Individual- und Klassenebene und mit Elterneinschätzungen ins Verhältnis gesetzt.

In *Studie 1* wurde der Frage nachgegangen, welche Überzeugungen Lehrkräfte zu Hochbegabung haben. Aus einer aktuellen Hochbegabungskonzeption wurden acht Dimensionen abgeleitet, die Überzeugungen zum Inhalt und zur Entwicklung von Hochbegabung erfragen sollten. Es wurden vier Gruppen miteinander verglichen: Lehramtsstudierende, Grundschullehrkräfte, und Kursleiterinnen und Kursleiter eines Begabtenförderprogram, die entweder schulische Lehrkräfte waren oder nicht. Zusätzlich wurden die Überzeugungen mit der Anzahl an Jahren im Lehrerberuf und im Begabtenförderprogram sowie mit Überzeugungen zur Veränderbarkeit von Intelligenz in Beziehung gesetzt. Die theoretisch abgeleiteten Überzeugungsdimensionen zeigten sich empirisch. Trotz einer großen Vielzahl an Überzeugungen, stimmten sie tendenziell mit der genutzten Hochbegabungskonzeption überein. Zwischen den Gruppen wurde partiell starke Messinvarianz erreicht. Die Überzeugungen der Lehramtsstudierenden unterschieden sich am stärksten von denen der anderen Lehrkräfte, während sich die Letzteren ähnelten. Keine oder wenige Unterschiede in den Hochbegabungsüberzeugungen zeigten sich anhand der Jahre im Lehrerberuf oder im Begabtenförderprogram. Jedoch waren die Hochbegabungsüberzeugungen mit den Intelligenzüberzeugungen assoziiert.

Studie 2 beschäftigte sich mit der Frage, welche Grundschulkinder von Lehrkräften für ein Begabtenförderprogramm nominiert werden. Es wurde angenommen, dass die Wahrscheinlichkeit nominiert zu werden positiv mit der Intelligenz der einzelnen Kinder zusammenhängt, sich jedoch das durchschnittliche Intelligenzniveau einer Klasse im Sinne eines Referenzgruppeneffektes negativ auswirkt. Zusätzlich wurde untersucht, welchen Einfluss Hochbegabungserfahrungen, Überzeugungen, dass Hochbegabung holistisch oder domänen-spezifisch ist, und Überzeugungen zur Veränderbarkeit von Intelligenz auf den Referenzgruppeneffekt haben. Wie erwartet waren die individuellen Intelligenzwerte positiv mit der Nominierungswahrscheinlichkeit assoziiert. Zudem zeigte sich der erwartete negative Referenzgruppeneffekt: Kinder hatten höhere Nominierungschancen in Klassen mit niedrigeren durchschnittlichen Intelligenzwerten als Kinder mit gleicher Intelligenz in Klassen mit höheren durchschnittlichen Intelligenzwerten. Der negative Referenzgruppeneffekt war besonders bei Lehrkräften deutlich, die Hochbegabung als holistisch anstatt als domänen-spezifisch ansahen. Die Intelligenzüberzeugungen hatten keinen Effekt, jedoch hingen verschiedene Arten von Hochbegabungserfahrungen unterschiedlich mit dem Referenzgruppeneffekt zusammen.

Mit *Studie 3* wurden zum einen Lehrereinschätzungen von nominierten Kindern mit Elterneinschätzungen verglichen und zum anderen der Einfluss der Kongruenz zwischen den beiden Ratern auf die Deutsch- und Mathematiknoten der Schülerinnen und Schüler untersucht. Erfasst wurden Einschätzungen zu verbalen und mathematischen Fähigkeiten, zum deduktiven Schlussfolgern, kreativen Denken und Engagement. Lehrkräfte und Eltern wurden in Bezug auf die Akkuratheit der Einschätzungen sowie den Einfluss von Halo-Effekten auf ihre Einschätzungen verglichen. Die Übereinstimmung in ihren Einschätzungen wurde ermittelt. Die Lehrer- und Elterneinschätzungen unterschieden sich nicht in ihrer Akkuratheit. Beide Rater waren von Halo-Effekten betroffen, Lehrkräfte jedoch stärker als Eltern. Die Übereinstimmung zwischen Lehrer- und Elternratings war niedrig bis mittelhoch. Positive Lehrer- und Elterneinschätzungen waren in additiver Weise mit besseren Deutschnoten verbunden. Positive Elterneinschätzungen der mathematischen Fähigkeiten und des Engagements reduzierten den Zusammenhang zwischen Lehrereinschätzungen und Mathematiknoten.

Die Ergebnisse der drei Studien werden in der abschließenden Diskussion zusammengefasst und kritisch beleuchtet. Des Weiteren werden Schlussfolgerungen für zukünftige Forschungsvorhaben und für die pädagogische Praxis abgeleitet.

CONTENT

1

Introduction and Theoretical Framework

1 Introduction and Theoretical Framework

The interest in gifted students and the willingness to support their giftedness has undergone great change in Germany (Heinbokel, 2001): Until the end of the seventies, giftedness was not a topic of discussion for the public, research, or schools. However, with the founding of the *Deutsche Gesellschaft für das hochbegabte Kind e.V.* (German Association for the Gifted Child, DGhK e.V.) in Hamburg and the establishment of gifted classes at the school *Christophorusschule* in Brunswick at the end of the seventies, giftedness began to attract attention. A breakthrough was the 6th World Conference of the *World Council for Gifted and Talented Children* (WCGTC) in 1985 in Hamburg. An intense political debate was started about whether gifted education programs are linked to elitism and, therefore, should not be supported in a democracy. On the other hand, discussions about how to help and nurture gifted children emerged. Nowadays, the identification and promotion of gifted students receives great interest in the public, politics, and research (Fischer & Müller, 2014). In particular, the goal of individualizing learning processes for all—including students who are high-achieving or show potential—by means of curricular and extracurricular interventions has been stressed in Germany (Fischer & Müller, 2014; Kultusministerkonferenz, 2015), but also, for example, in the US (Plucker & Callahan, 2014).

Many scholars have argued that gifted students need to be identified as early as possible to optimize their socialization and education processes (Fatouros, 1986; Heller, 2004; Karnes & Johnson, 1990; Schofield & Hotulainen, 2004). Elementary school teachers, therefore, play an important role in the development of gifted students. They are, for example, often involved in identification procedures for gifted education programs (National Association of Gifted Children, 2013), but also teach gifted students in general classrooms, which at the elementary school level are not separated by achievement in Germany. In the earliest school years, formal testing for giftedness is less likely to occur than at later stages (Gross, 1999), and, therefore, Brighton, Moon, Jarvis, & Hockett (2007) emphasized the significant role of elementary school teachers' beliefs about giftedness in the support that students receive to develop their talents.

The present dissertation addresses elementary school teachers' judgments of giftedness in students. It aims to provide a comprehensive view by discussing these judgments using a heuristic framework that is closely oriented on Südkamp, Kaiser, and

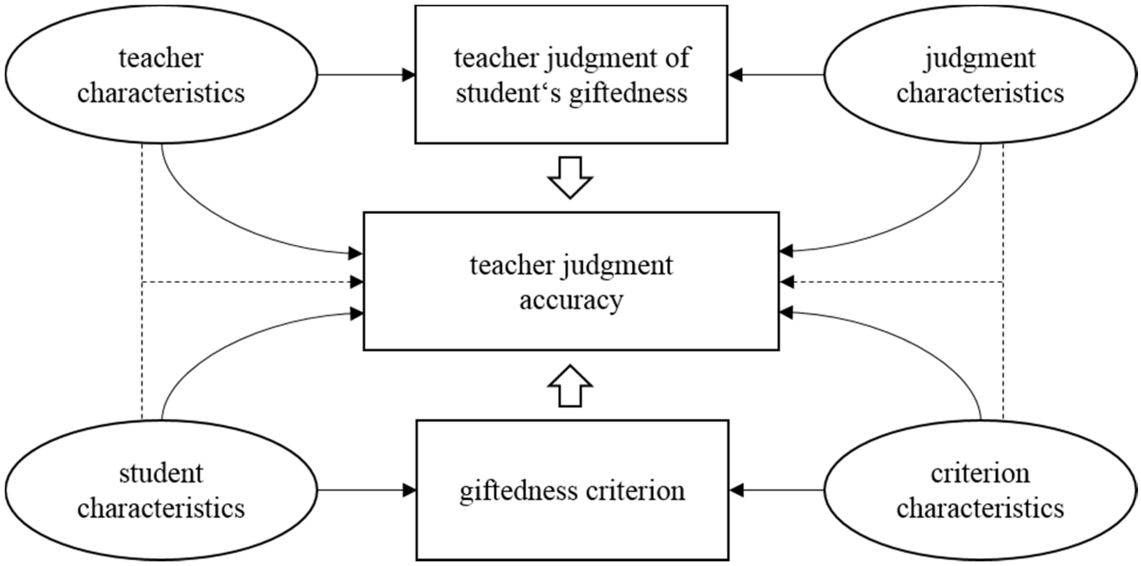

Figure 1.1. Heuristic framework of teacher judgments of students' giftedness, adapted from Südkamp et al. (2012).

Möller's (2012, see Figure 1.1) *model of judgment accuracy*. It systematizes moderators of teacher judgments and their accuracy into the following categories: (a) teacher variables like their beliefs about giftedness and experiences in the area of giftedness; (b) student characteristics such as students' intelligence on the individual and class level; (c) characteristics of the judgments like a global dichotomous judgment about giftedness or specific ratings of facets of giftedness on rating scales with more than two response categories; (d) the giftedness criterion, which might be based on a conception of giftedness and its operationalization into, for example, specific student tests; and (e) different kinds of measurements of the accuracy of teacher judgments such as correlations. Whereas a test of the whole model was beyond the scope of this dissertation, three empirical studies were conducted to focus on relevant teacher and student characteristics of teacher judgments with respect to students' giftedness.

First, in order to understand teacher judgments, teacher beliefs must be a focus of assessment as they can filter information and can guide interpretations and judgments (Five & Buehl, 2012; Shavelson & Stern, 1981). Systematic inquiries into teacher beliefs that align with central aspects of scientific conceptions of giftedness are rare in the relevant body of research (for an exception, see Schroth & Helfer, 2009). Hence, a comprehensive questionnaire was developed in Study 1 to measure beliefs about the content and development of giftedness that is based on a modern scientific conception of giftedness. The questionnaire was used on groups of teachers with different amounts of experience in general classrooms and in teaching gifted students. Furthermore, Study 1 examined how teachers' beliefs about the changeability of intelligence and the length of their tenure in general classrooms and in an enrichment program for gifted students were connected to beliefs about the nature of giftedness.

Next, teacher judgments of giftedness are mostly global dichotomous decisions of whether or not a student is gifted. In Study 2, these judgments are viewed on the basis of an investigation of teacher nominations of students for an enrichment program for gifted elementary school students. Nomination decisions were related to student variables like intelligence on an individual and class level. Specifically, the class-average level of intelligence has been rarely considered up to this point, although it has been found to be connected to elementary school teachers' judgments of cognitive abilities (Baudson, Fischbach, & Preckel, 2014). Furthermore, the effects of teachers' experiences in the area of giftedness, their beliefs on whether giftedness is holistic or domain-specific, and their beliefs on whether intelligence is malleable or fixed on their nomination decisions and the possible connection between the class-average level of intelligence and nominations were explored.

Lastly, despite the dichotomous nature of giftedness judgments, giftedness is typically seen as a combination of several characteristics like intelligence, creativity, and motivation (see Sternberg & Davidson, 2005). Although teachers might judge certain students to be gifted, students' levels in these facets of giftedness likely differ and might be rated differently by teachers and parents. Furthermore, the congruence between teacher and parent ratings might be relevant for students' academic development (Brenner & Mistry, 2007; Glueck & Reschly, 2014; Peet, Powell, & O'Donnel, 1997). Therefore, in Study 3, teacher and parent ratings of teacher-nominated gifted students' verbal and mathematical abilities, deductive reasoning, creative thinking, and engagement were compared with respect to their accuracy levels and the influence of halo effects. The congruence of teacher and parent ratings for each facet of giftedness was studied and connected to students' school grades.

The structure of this dissertation is as follows: The introductory chapter starts with a presentation of the scientific meaning of giftedness (Section 1.1), while Section 1.2 discusses some of the consequences of students being identified as gifted. Section 1.3 provides an in-depth analysis of teacher judgments and beliefs about giftedness. The accuracy of teacher judgments of students' giftedness will be presented. Teachers' beliefs about giftedness, further teacher variables, and student characteristics on the class level are also discussed. Furthermore, the accuracy of teacher judgments about academic achievement, cognitive ability, creativity, and motivation will be considered and compared with parent judgments. As the combination of these characteristics is important for judgments of giftedness, research on halo effects will be included. The introduction concludes with the research questions of the present dissertation (Section 1.4). The subsequent three chapters (Chapters 2-4) contain the three empirical studies that have been conducted. In the final chapter (Chapter 5), the main results of the three empirical studies will be summarized and discussed (Section 5.1). In Section 5.2, strengths and limitations of the dissertation will be outlined. Implications for future research and educational practice will be discussed in Sections 5.3 and 5.4. A final conclusion will be given in Section 5.5.

1.1 Scientific Conceptions of Giftedness

In this section, scientific conceptions of giftedness are discussed as the basis for an understanding of teachers' beliefs and judgments about giftedness. There are around 100 scientific answers to the question of what giftedness is (Freeman, 2005). A first approach to understanding why there is such a huge number of definitions of giftedness is the statement by Petersen (2013, p. 347): "Defining giftedness is a difficult challenge because the definition of high-ability varies across contexts and according to the values of each culture." This difficulty has been detected in similar ways by many researchers (Mayer, 2005), resulting in several notions: that there is and probably will never be one conception of giftedness that all can agree upon (Mönks & Katzko, 2005; Petersen, 2013), that conceptions to unite the field are needed (Mayer, 2005; Subotnik, Olszewski-Kubilius, & Worrell, 2011, 2012), and that a paradigm change is needed (Ziegler, 2005). Some are in favor of dropping the concept of giftedness entirely (Borland, 2005; Peters, Matthews, McBee, & McCoach, 2014).

Giftedness is a social construct, not an entity like weight or diabetes (Pfeiffer, 2015). It receives its properties through "the give and take of social interaction, not through the slow accretion of empirical facts about a preexisting entity, at least not exclusively" (Borland, 1997, p. 7). The same student might be gifted in Germany but not in China, based on one definition of giftedness but not on another, today but not tomorrow.

The term giftedness has been applied to different domains (Preckel & Vock, 2013). The Marland Report (1972, as cited in Ross et al., 1993, p. 23) that outlined the status of gifted education provision in the US identified six areas of giftedness: "1. general intellectual abilities, 2. specific academic aptitude, 3. creative or productive thinking, 4. leadership ability, 5. visual and performing arts, 6. psychomotor ability". The German word for giftedness is *Begabung* or *Hochbegabung* (Mönks & Katzko, 2005). Ziegler, Stoeger, Harder, and Balestrini (2013) examined the Deutscher Referenzkorpus (German Reference Corpus) and found that the term *Begabung* was strongly associated with the fine, liberal, and performing arts like music and acting, whereas *Hochbegabung* was typically associated with intellectual and academic giftedness. The focus of the present dissertation is on intellectual and academic giftedness. To analyze the differing meanings and conceptualizations of intellectual and academic giftedness, a presentation of the core themes along which conceptions differ will first be provided. As reactions to the current state of the field on conceptions of giftedness, second, Subotnik et al.'s (2011, 2012) mega-model as one attempt to unify the body of conceptions, and, third, Ziegler's (2005) systematic approach as a call for a paradigm change will be presented and discussed.

1.1.1 Systematizations of conceptions of giftedness

Conceptions of giftedness differ in diverse ways, and many scholars have proposed systematizations of these conceptions or collected key issues surrounding the reasons for differences between conceptions (Baudson, 2016; Dai, 2009; Hoge & Cudmore, 1986; Kaufman & Sternberg, 2008; Mönks & Katzko, 2005; Rost & Schilling, 2010; Sternberg, Jarvin, & Grigorenko, 2011; Subotnik et al., 2011). For example, Preckel and Vock (2013) differentiated between two dimensions: (a) uni- vs. multidimensional definitions and (b) competence vs. performance definitions. Pfeiffer (2015) distinguished between (a) the traditional psychometric view, (b) multiple intelligences, (c) expert performance models, and (d) talent development models. Although these systematizations are organized differently and discuss a broad spectrum of tensions and categories, several similarities in content are observable.

1.1.1.1 Potential to achieve versus actual achievement as criterion for giftedness

A major source of the diversity of conceptions is the question of what the criterion or marker for giftedness should be. Preckel and Vock (2013), for example, distinguished between competence and performance models. Competence models see giftedness as the potential for high achievement (e.g., Mönks & Katzko, 2005; VanTassel-Baska, 2005). Persons do not need to convert their potential into superior achievement to receive the giftedness label. Performance models, in contrast, equate giftedness with exceptionally good performance in a domain like mathematics, sports, or music (e.g., Plucker & Barab, 2005; Ziegler, 2005). Here, persons need to show exceptionally good performance to be identified as gifted. Hence, for students whose levels of potential do not match their levels of achievement (i.e., under- or overachievement), whether or not they receive the giftedness label is strongly dependent on the definition used. Preckel and Vock (2013) and Sternberg and Zhang (1995) noted that models that see giftedness as potential are mostly accepted for children, whereas performance definitions are normally used for adults. Other authors have also communicated the idea that giftedness criteria should be age-specific (Cross & Coleman, 2005; Subotnik et al., 2011).

The line between competence and performance definitions is not always clear. Following Mayer (2005), some scholars' definitions (e.g., Simonton, 2005; Brody & Stanley, 2005) can be classified under competence definitions as they state that extraordinary ability, talent, or capabilities are the criteria for giftedness, but propose using achievement tests to measure their criteria. Mayer noted that a solution might be to see giftedness as "developing or learning at a faster rate than one's cohort" (p. 440), which would allow the use of achievement tests as indicators for potential.

1.1.1.2 General and/or domain-specific intelligence

In the area of intellectual and academic giftedness, the role of intelligence is a main separator of giftedness models in two ways: Models can be differentiated in their conceptualization of intelligence and in the weight that intelligence has in the model. The development of giftedness models is closely related to the development of intelligence models (Borland, 1997; Kaufman & Sternberg, 2008): Traditionally, conceptions of giftedness focused on general intelligence (e.g., Hollingworth, 1942; Rost & Schilling, 2010; Terman, 1925). Later conceptions included the multidimensionality of intelligence (e.g., Brody & Stanley, 2005; Thurstone, 1938). Gardner (1983) broadened the view of giftedness with his model of multiple intelligences. He identified up to nine intelligences (e.g., interpersonal intelligence or mathematical-logical intelligence) that were seen as uncorrelated. His model was very influential in the area of giftedness but has been criticized, for example, for its insufficient empirical support (Kaufman & Sternberg, 2005; Pfeiffer, 2015; Rost, 2008). The hierarchical conceptions of intelligence—like the three-stratum model (Carroll, 1993) or the Cattell–Horn–Carroll (CHC) Theory of Intelligence (McGrew, 2009)—are particularly important for the present understanding of giftedness. In the CHC theory, general intelligence ("g") is at the highest level of the hierarchy, followed by more specific abilities like fluid reasoning (Gf) or comprehensive knowledge (Gc) on the next level that still apply to a broad range of domains. On the lowest level are diverse narrow abilities like lexical knowledge or perceptual speed. Many giftedness conceptions have incorporated the idea of an interplay between general and domain-specific abilities in explaining high competence or achievement in a domain (Kaufman & Sternberg, 2008; Subotnik et al., 2011).

1.1.1.3 Holistic versus domain-specific giftedness

A shift from a domain-general to a domain-specific view of giftedness was linked to the incorporation of domain-specific intelligence into conceptions of giftedness (Dai, 2009; Subotnik et al., 2011). A domain-general or holistic understanding of giftedness means that giftedness is seen as a general potential to develop high competencies across a wide array of domains. Although gifted persons will normally specialize in one or a few domains, their abilities and competencies can be flexibly used to be successful in other domains, too, if they so choose. However, most scholars have a domain-specific view of giftedness (Mayer, 2005). Giftedness is understood as high potential or excellent performance in a specific domain like mathematics or languages. Although persons might be gifted in more than one domain, their giftedness is mainly seen as a context-dependent fit and development of their (domain-specific but often also general) characteristics in a certain domain, with its unique conditions.

1.1.1.4 Uni- versus multidimensional view of giftedness

The second way in which the role of intelligence divides the field of giftedness is the weight that it receives in a model. Again, from a traditional perspective, giftedness is equated with intelligence. However, this unidimensional view has been criticized as too narrow if the goal is to explain (the development of) excellence (Borland, 2005; Preckel & Vock, 2013). Consequently, many multidimensional conceptions have been proposed that advocate including further cognitive characteristics like creativity and other intrapersonal characteristics like motivation to explain the development of excellence. For instance, Renzulli's (2005a) *Three-Ring Conception of Giftedness* proposed that giftedness in various domains can be developed at the intersection of well above average ability, task commitment, and creativity. He separated schoolhouse giftedness from creative-productive giftedness. The former refers to students with high ability or achievement test scores, the latter to students with innovative ideas and high-level productivity. With this separation, he addressed the tension of whether expertise or creativity is the essence of giftedness (Dai, 2009; Kaufman & Sternberg, 2008; Rost & Schilling, 2010).

1.1.1.5 The inclusion of environmental variables

Most of these multidimensional models also included students' environments. For example, the *Multifactor Model* (Mönks & Katzko, 2005) expanded Renzulli's (2005a) *Three-Ring Conception of Giftedness* by adding the environmental factors family, school and peers. The *Munich Model of Giftedness* (Heller, Perleth, & Lim, 2005) and the *Differentiated Model of Giftedness and Talent* (Gagné, 2005) also added these environmental factors as well as, for example, (critical) incidents. Furthermore, they differentiated more clearly between potential and performance than Renzulli's model and proposed interactive rather than additive relationships between the factors (Preckel & Vock, 2013). However, problematic issues surrounding the multidimensional conceptions mentioned are, among others, that not all variables are sufficiently theoretically described, empirically sound, and measurable and that as a result of the partially insufficient specification of the relations, the models are not empirically testable as a whole (Davidson, 2009; Preckel & Vock, 2013).

1.1.1.6 Giftedness as fixed or mutable

With the inclusion of non-cognitive and environmental factors, the developmental character of giftedness was introduced (Mayer, 2005). In this view, giftedness is seen as a development that occurs through interaction between a person and her or his environment, meaning that giftedness is relative to the context and changeable (Dai, 2009). Coleman and Cross (2005), for example, emphasized the necessity for students to actively engage in the process of transforming their high abilities into actual superior achievement. Ericsson (2014) considered deliberate practice in a domain and the number of opportunities to practice as crucial for exceptional performance. The developmental view stands in opposition to the view that giftedness is a stable trait of a person that normally will surface without special support. Based on this view, the main reasons for excellent performance lie within the gifted person him- or herself. For example, Robinson, Zigler, and Gallagher (2000) stressed the importance of gifted students' unique traits and superior cognitive abilities, and Geake (2009) saw giftedness as the result of brain structure and biological differences.

1.1.1.7 Comparison with peers

The determination that someone is gifted is oriented on social reference standards (Freeman, 2005; Ziegler, 2005). Gifted students are at the upper end of the distribution of certain abilities or types of performance in a population and, therefore, have higher ability or achievement levels than most others. This comparison is often made with peers (e.g., Renzulli, 2005a) or with other high-functioning persons in a domain (Subotnik et al., 2011). Lohman and Gambrell (2012) argued for using local instead of national norms. If giftedness is seen as faster and better learning, comparisons have to be made among students with similar learning opportunities (e.g., to the school class, to students who also have another native language, or to the school). Moreover, the proposed ranges of gifted persons or gifted achievements in a society vary from the upper 3% (e.g., Robinson, 2005; Brody & Stanley, 2005) to the upper 20% (Renzulli, 2005a). Mayer (2005) suggested that the identification of the upper 5% of a cohort as a compromise. Gagné (1998) suggested five different levels of giftedness: the 10%, 1%, .1%, .01%, and .001% students at the top of a distribution are mildly, moderately, highly, exceptionally, and extremely gifted. However, if (and if so, where) the cutoff for dividing gifted and non-gifted students should be made is a highly controversial issue in the area of giftedness. There are no sound empirical or theoretical justifications for a precise gifted/not gifted cutoff (e.g., a student with an IQ score of 129 is not gifted, but one with a score of 130 is gifted), for the use of the same cutoff across all definitions of giftedness or across all domains of endeavor, or for any one of the above mentioned levels as the cutoff (e.g., Borland, 2009; Freeman, 2005; Ziegler, 2005).

1.1.1.8 Quantitative versus qualitative differences

There is also debate about whether quantitative or qualitative differences exist between gifted and non-gifted persons (Dai, 2009; Baudson, 2016). According to the quantitative differences view, gifted individuals differ in degree from their non-gifted peers: for example, in the speed or quality with which they acquire and process information. But some scholars assume further qualitative differences, arguing that gifted individuals also have information-processing abilities that differ in kind from their non-gifted peers (Keogh & MacMillan, 1996). The approach of categorizing students into gifted and non-gifted groups based on a certain cut-off score on a criterion is often motivated by statistical pragmatism but also communicates a view of qualitative, but not quantitative differences, treating the gifted as a homogeneous group despite evidence stressing the heterogeneity of gifted students (Dai, 2009; Baudson, 2016). Dai (2009) and Rogers (1986) highlighted a possible integration of both views: Differences between gifted and non-gifted students are of a quantitative nature. However, accumulated over time, they can lead to qualitative differences through, for example, different experiences. However, the debate about quantitative and/or qualitative differences remains unresolved (Keogh & MacMillan, 1996).

The presented issues surrounding conceptions of giftedness stress how differently giftedness has been conceptualized. Peters et al. (2014, p. 3) stated that "[w]hat can be said most confidently about conceptual definitions of giftedness, talent, and high ability is that they are widely inconsistent." Subotnik et al. (2011, 2012) developed a mega-model of giftedness in an attempt to unify the field. It presents an aggregation of the features that were assumed to be agreed upon among most scholars. However, Ziegler (2005) calls for a paradigm change in the conception of giftedness. His system view of giftedness is an important supplement to the presented debates about tensions between models, offering an alternative proposal for practice and research surrounding the giftedness construct. These two conceptions—Subotnik et al.'s (2011, 2012) *Talent-Development Mega-Model* and Ziegler's (2005) *Actiotope Model of Giftedness*—are discussed in more detail in the following sections.

1.1.2 An attempt at unification: The Talent-Development Mega-Model

Subotnik et al. (2011, 2012) provided a comprehensive definition based on a seminal review of the giftedness literature that has gained broad attention as potentially providing a common foundation for giftedness definitions (for comments and criticisms, see, e.g., Grantham, 2012; Worrell, Olszewski-Kubilius, & Subotnik, 2012; Ziegler, Stoeger, & Vialle, 2012). Subotnik et al.'s definition covers all domains of possible giftedness (e.g., academics, music, and sports) and combines the main characteristics of modern conceptualizations of giftedness such as the *Three-Ring Conception of Giftedness* (Renzulli, 2005a), the *Munich Model of Giftedness* (Heller, Perleth, et al., 2005), and the

Differentiated Model of Giftedness and Talent (DMGT; Gagné, 2005). At the same time, it distances itself from more traditional conceptualizations (e.g., Hollingworth, 1942; Rost & Schilling, 2010; Terman, 1954) that see giftedness as more stable and mainly limited to intelligence.

Subotnik et al.'s (2011, 2012) conception will be presented in alignment with the discussion of the main differences between conceptions of giftedness presented above (see Section 1.1.1 and Table 1.1, p.16). First, Subotnik et al. offer a solution to the question of whether giftedness is the potential for, or actually exhibited superior performance. The authors see both as criteria for giftedness but at different phases of the development of giftedness. Whereas high potential is a reasonable criterion at the beginning stage of giftedness, superior achievement has to be shown at later stages. In this conception, a third criterion for persons with fully developed giftedness is introduced: eminence. Eminence means that persons have to be recognized as one of the top experts in their field. Specifically, the eminence criterion emphasizes that giftedness is seen as something that a society grants someone and not as a trait that a person possesses.

Second, Subotnik et al. (2011) see high intelligence as a necessary condition for explaining giftedness in intellectual and academic domains. Both general and domain-specific abilities are important; however, the role of general intelligence differs according to the domain of the gift, and the importance of domain-specific abilities increases with a person's progress in a domain. Third, Subotnik et al. point out that the start, peak, and end of giftedness are strongly tied to the corresponding domain. For example, a person who is gifted in mental arithmetic has a different course of development than a person who is gifted in the area of psychotherapy. They will show their potential, achievement, and eminence at different points in their lives and for periods of time differing in length. Fourth and fifth, high intelligence is not sufficient for explaining intellectual and academic giftedness. Psychosocial skills, motivation, and environmental variables like opportunities are also needed to translate potential into high achievement. The ratio and interactions between the involved variables, however, are not specified for the various domains. Sixth, giftedness is seen as a malleable and developmental construct that results from and develops further via an interplay of biological, pedagogical, psychological and psychosocial factors. To nurture giftedness, focus needs to be directed not only on the development of one's ability, but also on other person-related and environmental variables.

Seventh, gifted individuals need to have higher potential or achievement than most other highly able persons in a domain. The authors did not define clear cut-offs but did state that more individuals will be recognized as gifted in the beginning phase of a giftedness (when the key variable is potential) than in the last stage in which a person has to reach eminence. Eighth, Subotnik et al. (2011) see quantitative and qualitative differences between gifted and non-gifted persons and stated that “gifted individuals are different by virtue of their combination of intensity, persistence, and ability that results in eminent productivity” (p. 21). The heterogeneity of giftedness manifestation and, by implication, of gifted individuals is stressed.

1.1.3 A systematic perspective on giftedness: The Actiotope Model of Giftedness

Ziegler (2005) criticized the relative lack of learning and environment orientation in conceptions of giftedness and, hence, in the diagnosis and promotion of giftedness. In a call for a paradigm change, Ziegler et al. (2012) noted that “it is not enough to pay lip service to the importance of the environment or to fragment the research field into gifts (talents, abilities, etc.), internal moderators (e.g., high motivation), and external moderators (e.g., mentors), which collude in a simple summative or multiplicative manner. Better suited are models within the ecological or systemic paradigm…” (p. 196). Ziegler (2005) proposed a systematic model of giftedness—the *Actiotope Model of Giftedness*—that focuses on the interactions between individuals and the specific systems that surround them to explain the development of excellence. For the development of excellence, there has to be an effective interplay between the person and her or his environment, goals, and learning. Hence, Ziegler argued that research on giftedness has to move away from an investigation of individual personality traits to an examination of the system surrounding an individual and her or his environment, which leads to specific actions.

Before stating the model’s positions concerning the main categories of differences among conceptions of giftedness (see Section 1.1.1), the main assumptions of the *Actiotope Model of Giftedness* will be described. The core of the model is the *actiotope* that comprises the person and the environment with which she or he can interact. More concretely, four components have to be differentiated: (A) The *action repertoire* of a person consists of all actions that a person has objectively at her or his disposal in a certain situation. People’s action repertoires differ on the basis of inter- and intrapersonal determinants. A fifth-grader will probably have a smaller action repertoire in physics than a tenth-grader, for example, based on school curriculum. Intrapersonal determinants like ability are at the core of many conceptions of giftedness. Therefore, these models can be seen as subtheories of the Actiotope Model of Giftedness. (B) A persons' *goals* contribute to actions. Particularly, goals to develop excellence and to exercise excellent actions are central for giftedness. (C) The *environment* that frames someone's actions has an important influence on the development of excellence, including, for example, a person's family, peers, and teachers, the resources that a person can draw on, and the specific setting of the talent domain that dictates what actions are seen as excellent. (D) The *subjective action space* is a person's internal representation of all possible actions that can be conducted in a situation with certain conditions. A person chooses a specific action based on this representation. The subjective action space does not have to be congruent with the objective action repertoire. A person might misjudge which actions she or he can perform.

As situations change, these four components have to coevolve and the interactions between them have to adapt effectively in order for a person to achieve excellence. Persons need to learn, for example, to distinguish between successful and unsuccessful actions in certain situations, to recognize conditions for the execution of actions, to generate variations of actions, and to anticipate the usage of knowledge and actions. Feedback and discussion of possible future actions under certain conditions are needed for effective adaptation. Furthermore, based on the complex and permanent changes of the actiotope, the system needs to be stable to ensure effective development to excellence. An actiotope is stable if its components are co-adaptive and complementary (Ziegler & Philipson, 2012).

Hence, with regard to the categories of main differences between conceptions of giftedness (see Table 1.1, p.16), first, outstanding actions are defined as excellent. Persons who possibly and probably manifest excellence are seen as talented or gifted. Second, the role of intelligence is not the focus of the model; instead, mastery of preceding learning steps is crucial for the development of excellence. Third, giftedness is seen as domain-specific, and, fourth, as multidimensional. Fifth and sixth, the model includes environmental variables and sees giftedness as malleable. Seventh, the assessment of what actions are outstanding is not based on a general social norm across all domains (e.g., the 5 percent of most excellent individuals in a domain) but on a comparison with

Table 1.1

Description of Conceptions of Giftedness in Relation to Categories of Main Differences Between Conceptions of Giftedness

Categories	Talent-Development Mega-Model (Subotnik et al., 2011, 2012)	Actiotope Model of Giftedness (Ziegler, 2005)
1. Potential or achievement as criterion for giftedness	Potential (start), achievement (later), eminence (end)	Possibly realize excellence (talented), probably realize excellence (gifted), already realized excellence (excellent)
2. General and/or domain-specific intelligence	Both	Not the focus of the model (focus on mastery of learning steps)
3. Holistic vs. domain-specific giftedness	Domain-specific giftedness	Domain-specific giftedness
4. Uni- vs. multi-dimensional giftedness	Multidimensional view	Multidimensional view
5. Inclusion of environmental variables	Yes	Yes
6. Giftedness as fixed or mutable	Mutable	Mutable
7. Comparison with peers	With high-functioning individuals in a domain	Comparison with current achievement level in a domain
8. Quantitative versus qualitative differences	Both	Actiotopes are qualitatively different

current achievement levels in a domain. Actions of individuals who are assessed as possibly or probably able to manifest excellence receive the labels talented or gifted. Eight, systems are of concern in this model, and Dai (2009) indicated that the differences in the actiotope are of a qualitative nature.

Overall, there is no uniform definition of giftedness. Different conceptions of giftedness can lead to the identification of different students as gifted and to different explanations for their giftedness (see Sternberg and Davidson, 2005). The work on systematizations to extract key characteristics from groups of conceptions and discuss the main tensions between them can provide a framework of what can be subsumed under the construct of giftedness and in which direction the development of conceptions might go. Whether the steps taken by scholars like Subotnik et al. (2011) and Ziegler (2005) will lead to a unification of the field remains to be seen. For now, there is no one definition of giftedness that teachers should know or with which they are confronted in the school setting. This has to be kept in mind in discussions of teachers' beliefs about giftedness and their judgments about giftedness in students. Moreover, how teachers' beliefs about giftedness are linked to scientific conceptions of giftedness is discussed and empirically investigated in this dissertation.

1.2 Consequences of Teacher Judgments of Students' Giftedness

Several reasons can be given why teachers' beliefs and judgments about giftedness are relevant for students' development and therefore should be investigated. Wild (1993) summarized four reasons why teachers might have to judge whether or not a student is gifted: Their ratings are used as selection instruments for gifted education provision, for ability diagnoses within the general classroom, for individual counseling, and to form study samples for giftedness research. The following sections address the first two reasons: (1) teachers' role in the identification of students for gifted education and the consequences of (not) receiving gifted education, and (2) possible consequences of teachers' judgments of ability and achievement in the general classroom. Furthermore, (3) consequences of congruence between teachers' and parents' judgments of ability are outlined.

1.2.1 Gifted education

Teachers are frequently involved in the process of identifying students for gifted education services (Coleman & Gallagher, 1995; Schack & Starko, 1990; National Association of Gifted Children, 2013). Scholars (Heller & Perleth, 2008; Jarosewich, Pfeiffer, & Morris, 2002; Renzulli, 2005a) have recommended their inclusion because they can observe students in diverse learning and achievement-related situations, have a professional educational background, and can compare students against a broad reference group of other students in a class or school. If involved, they either initiate the giftedness assessment with their nominations of possibly gifted students, provide information about students during a multimethod assessment, or both (McBee, Peters, & Miller, 2016; McClain & Pfeiffer, 2012; Renzulli, 2005a). In praxis, teacher judgments are given great weight in decisions about interventions (Hoge, 1983), and sometimes are the sole selection method for gifted education programs (e.g., Deku, 2013; Freeman & Josepsson, 2002). Using teacher nominations as a screening instrument for potentially gifted students is often seen as pragmatic. It limits the number of students who undergo a more thorough assessment that might be costly and time-consuming. Putting teachers in this position of responsibility as "gatekeepers" has been recently criticized, especially due to the implication that students who are not nominated will not have a chance to be assessed (Acar, Sen, & Cayirdag, 2016; McBee et al., 2016).

1.2.1.1 Consequences for students' academic development

Overall, many but not all studies illustrate positive effects of gifted education on students' academic achievement. Wai, Lubinski, Benbow, and Steiger (2010) could show with two longitudinal studies among persons who were gifted in science, technology, engineering, and mathematics (STEM) that participation in diverse advanced and intellectually challenging educational interventions during one's school years is associated with higher academic achievement as an adult. Research on the effectiveness of specific kinds of gifted education have also been conducted. Gifted education can be separated into *acceleration* and *enrichment* or is a combination of the two (Fischer & Müller, 2014): First, acceleration is progress through the school system at a faster rate or at a younger age than same-age peers (Pressey, 1949, as cited by Southern & Jones). Examples are grade skipping and earlier entrance to school or university. Many but not all studies have shown positive effects of acceleration on students' academic achievement (Kretschmann, Vock, & Lüdtke, 2014; Kulik & Kulik, 1984; Steenbergen-Hu & Moon, 2011). Second, enrichment provides students with instruction or learning content in addition to the regular school curriculum, like extracurricular courses or visits to a museum. Enrichment can be vertical (i.e., more in-depth studies of topics in the curriculum) or horizontal (i.e., activities beyond the curriculum) and can take place during or outside of school time (Nogueira, 2006). Predominantly positive effects on students' academic achievement have been reported for enrichment (Kim, 2016; Vaughn, Feldhusen, & Asher, 1991). Furthermore, a meta-analysis by Kulik and Kulik (1992) on different kinds of ability grouping showed small effects for within-class grouping on students' academic achievement but stronger effects for gifted classes.

1.2.1.2 Consequences for students' social-emotional development

The consequences of participation in gifted education services on social-emotional development seem to be mostly positive for enrichment (Kim, 2016) and neutral to slightly positive for acceleration (Steenbergen-Hu & Moon, 2011). Specifically, the academic self-concept (i.e., students' subjective ratings of their own abilities in school generally or in specific academic areas; Trautwein, Lüdtke, Köller, & Baumert 2006) of gifted students in different educational settings has received great attention (Preckel & Vock, 2013) and is discussed under the terms of the *Big-Fish-Little-Pond Effect* (BFLPE; Marsh & Parker, 1984) and the *Basking-In-Reflected-Glory Effect (BIRGE*; Marsh, Kong,

& Hau, 2000). The BFLPE describes that students' academic self-concept is negatively associated with class-average or school-average ability levels because students perceive the contrast between themselves and others. Hence, a student who is in a class with a lower average level of ability will have a higher self-concept than a student with the same abilities who is in a class with a higher average level of ability. However, students might also perceive their similarity to a group, resulting in the assimilation effect BIRGE. For instance, a student can have a higher self-concept because she or he belongs to a higher-ability group than a same-ability student in a group with a lower average ability level. Marsh et al. (2000) indicated that both effects are active simultaneously but that the BFLPE is stronger than the BIRGE. Concerning gifted students, Preckel and Vock (2013) reported that gifted students' academic self-concept in gifted classes is mostly but not always higher than that of average-ability students in regular school classes. If students with the same ability level are compared, gifted classes seem to be associated with negative to neutral effects (Marsh, Chessor, Craven, & Roche, 1995; Preckel & Brüll, 2010; Zeidner & Schleyer, 1999) and summer programs with neutral to positive effects (Cunningham & Rinn, 2007; Makel, Lee, Olszewki-Kubilius, & Putallaz, 2012) on students' academic self-concept.

1.2.1.3 Consequences of misnomination

If gifted students are overlooked, they might not receive gifted education but face continually unchallenging situations, which have been linked to behavioral and social problems (Heller & Schofield, 2008). A review of over 25 years of phenomenological qualitative research on gifted students' experiences in school (Coleman, Micko, & Cross, 2015) summarized that gifted students in schools that did not support their strengths more often reported feeling that they were different from other students, that they had to wait in class, were intellectually unchallenged, and were victims of bullying. Also, as already mentioned, gifted students who participated in gifted education had higher achievement as adults than gifted students who did not receive gifted education (Wai et al., 2010). If teachers falsely nominate a student for gifted education, the consequences of students' failure can differ according to the kind of gifted education. McBee et al. (2016) noted that dropout from an out-of-school enrichment program might be associated with only minimal negative consequences, whereas in a case of unsuccessful grade skipping, reintegration into the original classroom might be logistically consuming and socially negative.

1.2.2 General classroom

Virtually all elementary school teachers will teach gifted students over the course of their careers because gifted students can be found in all age groups, classes, and regions (Bangel, Moon, & Capobianco, 2010; Ross, 1993). Teachers' competence in rating learning and achievement-related student characteristics like giftedness forms—in addition to their ability to assess the demands of learning activities and materials—part of teachers' diagnostic skills, which are needed to construct tests, to grade, and to monitor students (Brunner, Anders, Hachfeld, & Krauss, 2013). Südkamp et al. (2012) emphasized, aside from teachers' involvement in placement and intervention decisions, four consequences of teachers' judgments of ability and achievement: First, teachers partially shape their instructional decisions on the basis of their ratings of students' performance, for example, for their decisions on instructional materials and teaching strategies (Shavelson & Stern, 1981). Second, teacher judgments of students' abilities are often expressed in school grades and, therefore, provide feedback to students and their parents (Hoge & Coladarci, 1989) and are relevant for students' academic careers (Fischbach, Baudson, Preckel, Martin, & Brunner, 2013; Maaz, Trautwein, Lüdtke, & Baumert, 2008). Third, through school grades, teacher judgments affect students' academic self-concepts. Fourth, teacher judgments of students' abilities are associated with their expectations of students' abilities. Judgments seem to inform expectations but expectations about students' ability can also guide further judgments (Rubie-Davies, 2010; Timmermans, de Boer, & van der Werf, 2016). Teacher expectations are often correct and can predict future student achievement, but if they are inaccurate, they can also have, in the form of self-fulfilling prophecies, effects on students' future achievements—although these effects are mostly small (de Boer, Bosker, & van der Werf, 2010; Jussim & Harber, 2005). Therefore, an overestimation of students' abilities can have beneficial effects on students' academic achievements, although the judgment that a student is gifted can also lead to pressure (Moulton et al., 1998). An underestimation can lead to a climate that encourages a level of achievement that is below a student's potential (i.e., underachievement; Kolb & Jussim, 1994).

If students are openly labeled as gifted, several effects of this label have been reported for teachers, students, the labeled student, and their family members (Berlin, 2009; Hickey & Toth, 1990; Matthews, Ritchotte, & Jolly, 2014): Study participants saw—often simultaneously—positive consequences of the giftedness label like the opportunity to receive gifted education and to interact with other gifted students, and negative consequences like social damage through negative stereotypes of gifted students. Manaster, Chan, Watt, and Wiehe (1994) showed that if gifted students perceived being negatively stereotyped, they indicated that it came from classmates in general—students who they did not know well—not from persons whose opinions they valued like teachers, friends, or parents. This result is in line with contact theory, which states that intergroup contact reduces prejudice and stereotypes (Pettigrew & Tropp, 2006). The "stigma of giftedness" paradigm (Coleman & Cross, 1988; Cross, 2005) and the notion of stereotype threat in Matthew et al.'s (2014) review emphasize that some, but not all, students labeled as gifted react to anticipated negative stereotypes with highly confirming behavior (e.g., showing off), with hiding their abilities (e.g., not admitting that tests are easy), or with behavior to show that they do not identify with gifted students (e.g., befriending groups that are considered to be different from gifted students).

With regard to possible educational consequences of the giftedness label, Dweck (1999) was concerned that it transmits the impression that intelligence is fixed and cannot be changed. This view has been found to be associated with a focus on self-defense of one's abilities instead of self-improvement (Blackwell, Trzesniewski, & Dweck, 2007; Carr, Rattan, & Dweck, 2012). Also, concern has been expressed that the giftedness label can imply to teachers that students will make it on their own and do not need special support (Moon, 2009). Overall, scholars like Borland (2005), Peters et al. (2014), Renzulli (2005a), and Ziegler (2005) criticized the praxis of labeling persons as gifted and advocated a change towards seeing behavior, systems, or programs as gifted.

1.2.3 Being seen as gifted by teachers and parents

As outlined above, teachers play an important role in the identification and education of gifted students. However, teachers' judgments about students have to be seen in part in connection with students' parents, as the congruence between teacher and parent judgments can have effects on students (Glueck & Reschly, 2014). Two reviews (Christenson, 1999; Glueck & Reschly, 2014) on school-family relationships indicated that students can perform best when teachers and parents are congruent in realistically high standards and expectations, in opportunities to learn, in support and guidance, in the kind of climate and relationship (e.g., warmth, encouragement, and acceptance) towards a student, and in modeling desired behaviors and commitment to learning. Furthermore, high congruence in teacher and parent judgments of students' competence and engagement in school (Peet et al., 1997) and in their judgments of students' future school success (Brenner & Mistry, 2007) was connected to higher school achievement.

Peet et al. (1997) contemplated that similar or dissonant teacher-parent judgments of students' achievement might result in comparable or conflicting educational feedback or advice. Hoge and Kremp (2006) indicated that some parents and teachers mutually distrusted each other's abilities to judge students' giftedness and their educational responses to students. Teachers' mistrust of parents was seen as one possible reason in this study for why teachers did not recognize that students were gifted. Moreover, negative effects on teacher expectations of students' academic achievement have been reported if teachers perceived differences between their own values and parents' values on, for instance, appropriate child-caring or the significance of education (Brighton et al., 2007; Hauser-Cram, Sirin, & Stipek, 2003). However, Brenner and Mistry (2007) could show that high parent expectations could soften the effect of low teacher expectations on students' school achievement. Still, Glueck and Reschly (2014) indicated that empirical support for the connection between teacher-parent congruence and students' school achievement is sparse.

Overall, teachers' judgments about students' giftedness can be associated with diverse consequences for students' development. This dissertation therefore addresses these teacher judgments with three empirical studies by, first, exploring what teachers believe giftedness means, second, investigating teachers' nomination decisions for an enrichment program, and, third, examining teacher judgments of students' facets of giftedness, comparing them to parent judgments, and linking teacher-parent congruence in judgments to students' school grades.

1.3 Teacher Judgments and Beliefs about Giftedness

The previous section presented the important role that teacher judgments of ability in general and of giftedness specifically can play for students. This section will deal with teacher judgments of giftedness and discuss important factors that are connected to these judgments. Südkamp et al.'s (2012) *model of judgment accuracy* was adapted as a heuristic framework for viewing teacher judgments about students' giftedness in this dissertation (see Figure 1.1, p. 4). This systematization highlights the need to consider teacher, judgment, student, and criterion characteristics as well as different kinds of accuracy measurements in order to explore teacher judgments about giftedness. The following sections will address research on these characteristics.

Research on the accuracy of teacher judgments of students' giftedness is presented first, taking into consideration the kind of accuracy estimates and the characteristics of the criterion and judgment. Second, research on teachers' beliefs about giftedness as a significant teacher variable is reviewed and put in the context of scientific conceptions about giftedness. Further sections address third, the associations of further teacher variables, fourth, students' demographic variables, and fifth, facets of students' giftedness with teacher judgments and beliefs about giftedness. Sixth, the importance of considering reference group effects on teacher judgments about giftedness is highlighted and, seventh, teacher judgments are set into relation with parent judgments about giftedness.

1.3.1 Teachers' accuracy in judging students' giftedness

How accurately teachers can detect giftedness in students has been a frequent subject of research. Judgments about giftedness are mostly dichotomous decisions whether or not a student is gifted but can also be based on rating scales on which teachers rate facets of students' giftedness (Hoge & Cudmore, 1986). To estimate the accuracy of giftedness judgments, it has to be considered which giftedness criterion teacher judgments are tested against. As outlined in Section 1.2, some scholars see giftedness as high potential, others as actual superior performance. Furthermore, some researchers use intelligence as the sole characteristic for explaining giftedness, whereasothers combine several factors like intelligence, creativity, motivation, and environmental variables (Preckel & Vock, 2013). Additionally, different methods are used to measure these characteristics, for example, with tests, work samples, and ratings and nominations by teachers, school committees, parents, counselors, and peers—as single methods or in combination (Carman, 2013; Coleman & Galagher, 1995; National Association of Gifted Children, 2013; Ziegler & Raul, 2000). Some scholars like Renzulli and Delcourt (1986) have emphasized that students' success in a gifted education program should be a criterion for the quality of teacher nominations. Based on this diversity, it is difficult to determine the "true" accuracy of teacher nominations (e.g., Hoge & Cudmore, 1986; McBee, 2006).

Moreover, different kinds of measures are used to estimate different aspects of the accuracy of judgments. Generally, three components of accuracy can be differentiated (Schrader & Helmke, 1987): the *level component* that allows statements about whether teachers over-, correctly or underestimated students, the *differentiation component* that indicates, for example, whether teachers over-, correctly or underestimated the variance in cognitive ability among students in a class, and the *correlational component* that shows, for example, how accurately teachers can sort students' cognitive abilities into a rank order. In giftedness research, effectivity and efficiency measurements are often reported (see Acar et al., 2016). Effectivity is the percentage of students who are nominated by teachers and are gifted (e.g., as determined by an IQ score above 130) in relation to all gifted students in a sample. Efficiency is the percentage of gifted students in the group of nominated students. The use of effectivity-efficiency measures was seriously criticized by Hoge and Cudmore (1986). Although the measures are dependent on the base rates, base rates were rarely reported and statistical significance tests were missing. Furthermore, Gagné (1994) criticized that the two measurements are negatively correlated. He proposed using the phi (Φ) coefficient that contains the correlation of the two dichotomous variables *nomination status* and *giftedness criterion*.

A review of research results on the accuracy of teacher judgments indicates that teacher nominations seem to be more effective than efficient (Heller, Reimann, & Senfter, 2005; Neber, 2004; Wild, 1993). An often cited study is the one by Pegnato and Birch (1959) in which teacher nominations—in addition to, for example, honor roll listing and group intelligence and achievement tests with different cut-off levels—were compared with students' scores in an individually administered intelligence test. If students had an IQ score of 136 or higher, they were identified as gifted (6.5% in this study). Teachers overlooked more than half of the gifted students (i.e., 45% effectivity). Furthermore, nearly three out of four of the nominated students were not gifted (i.e., 27% efficiency). The authors concluded that teacher nominations should not be relied upon for the identification of gifted students. A review of 22 studies (Hoge & Cudmore, 1986) that included the study by Pegnato and Birch reported high variance for both measures: 0-86% effectivity and 4-78% efficiency. Higher values than in Pegnato and Birch's study but a similar ratio were reported in a recent meta-analysis (Acar et al., 2016) that examined two groups of methods for identifying gifted students. They distinguished non-performance methods like nominations by teachers, parents, and students themselves from performance methods that included tests of academic achievement, cognitive ability, and creativity. If the performance methods were used as giftedness criteria, nonperformance methods were interpreted as effective (59%) but less efficient (39%). Teacher nominations did not differ significantly in their effectivity or efficiency from the other nonperformance methods.

Gagné (1994) reanalyzed Pegnato and Birch's (1959) study with the phi coefficient. Based on these analyses, the accuracy of teacher nomination was with $\Phi = .29$ not significantly lower than the accuracy of the group achievement test and only in two out of four cases significantly lower than the group intelligence tests (if the cutoff IQ for giftedness was > 120 or > 125, but not if the cutoff IQ was > 115 or > 130). Similarly, in the already mentioned meta-analysis by Acar et al. (2016), nonperformance methods correlated moderately with performance methods ($r = .30$). Heller, Reimann, et al. (2005) concluded based on correlations with an ability test that elementary school teachers were satisfactorily to sufficiently accurate in identifying gifted students. A meta-analysis by Machts, Kaiser, Schmidt, and Möller (2016) compared, among other things, teacher giftedness nominations with intelligence tests and reported correlations that ranged from $r = .04$ to $r = .52$ across the seven included studies, again with a moderate mean correlation. Wild (1993) controlled for the fact that several students were rated by the same teacher und noted that the spectrum of correlations between teacher judgments and an intelligence test varied strongly from nearly perfect congruence to zero correlations between judgments and tests. If a more comprehensive giftedness criterion that included tests of achievement, cognitive ability, creativity, and motivation was used, high accuracy was noted in a study by McBee (2006). Teacher nominations of students as gifted were with $\Phi = .51$ highly accurate but were outperformed by standardized tests ($\Phi = .68$).

If students' school success is viewed as criterion, Foreman and Gubbins (2014) showed that students whose teachers nominated them as having high learning potential benefitted more from a mathematics intervention than students who were not nominated. After controlling for pretest and reasoning ability scores, they received higher posttest scores. Also, Hunsaker, Finley, and Frank (1997) analyzed the success of students who were nominated by their teachers. Teacher judgments of facets of giftedness were positively associated with students' performance in a gifted education program as rated by the program's instructors.

Concerning different types of teacher judgments of giftedness, rating scales for judging facets of giftedness can be an effective aid for teacher judgments about giftedness when they provide a focus on relevant and observable characteristics (see Westberg, 2012, for a review). In Acar et al.'s (2016) meta-analyses, teacher judgments' accuracy was higher if they used ratings scales instead of solely giving a dichotomous judgment about students' giftedness. Furthermore, Ashman and Vukelich (1983) reported that it is preferable to use rating scales with items that have more than two response categories. This result was also supported by Machts et al.'s (2016) meta-analysis.

Overall, teachers tend to misidentify more students as gifted than they overlook gifted students. They are on average moderately accurate in judging giftedness in students, but the level of accuracy measured differs greatly across studies and teachers. Higher accuracy levels are reported when the giftedness criterion was multifaceted, more closely linked to student achievement, and based on rating scales.

1.3.2 Beliefs about giftedness

In praxis, many teachers seem to rely strongly or even solely on their beliefs about giftedness because they seem to lack knowledge about giftedness. This is true, for example, in the US (Brighton et al., 2007; National Association of Gifted Children, 2013), but also in Germany. Gifted education has become an important issue in Germany (Kultusministerkonferenz, 2015). The education acts in 12 out of 16 German states explicitly include gifted education; however, only two states cover it in their teacher training provisions (Fischer, 2014). For interested teachers, several voluntary education offers are available. In addition to a number of conferences on gifted education like the "Münster Congress of Education" ("Münsterscher Bildungskongress"), some universities, for example those in Berlin, Karlsruhe, Leipzig, and Münster, provide special training programs in the area of giftedness (Fischer & Müller, 2014). However, over 80% of around 3,500 elementary school teachers in Bavaria who participated in a study by Heller, Reimann, et al. (2005) stated that they had not received sufficient information about giftedness at university or during their job. Less than 2% reported that they had had special courses about giftedness during their teacher training. Also, in a study by Schneider, Preckel, and Stumpf (2014), around one-third of the secondary school teachers who taught gifted classes in Baden-Württemberg and Bavaria reported that they had not received any form of training or preparation in gifted education. Hence, as most educational programs in the area of giftedness are not compulsory in Germany, scientific conceptions of giftedness are often not sufficiently covered in teachers' general teaching education.

Scientific conceptions of giftedness are explicit theories that have a theoretical and/or empirical underpinning, define their components and mechanism, and are internally consistent and testable (Davidson, 2009; Sternberg, 1985). In contrast, teachers' beliefs about giftedness are often based on experiences and (subjective) reports about giftedness and are not necessarily in line with empirical results (Baudson, 2016; Baudson & Preckel, 2013a; Sternberg, 1985; Watt & Richardson, 2014). Beliefs have also been researched under terms like implicit theories, conceptions, opinions, and attitudes. However, whether these terms are synonymous or not has not yet been clarified (Pajares, 1992). Beliefs can be defined as "psychologically held understandings and assumptions about phenomena or objects of the world that are felt to be true, have both implicit and explicit aspects, and influence people's interactions with the world" (Voss, Kleickmann, Kunter, & Hachfeld, 2013, p. 249). Beliefs are intertwined, interact with other beliefs in a complex and multifaceted belief system, and can have different functions. As filters, they shape teachers' perception and interpretation of information and experience; as frames, they are used to define problems and tasks; and as guides, they shape persons' goals and behaviors (Five & Buehl, 2012). Brunswik's lens model (1955; as cited by Kleber, 1992) indicates that in making judgments about a latent characteristic such as giftedness, observable indicators such as behavior, test scores, or biological characteristics are used, and that the usage and weighting of these indicators in making judgments are associated with various teacher variables such as their beliefs. Similarly, Shavelson and Stern (1972) stated in their heuristic framework of teacher judgments that student information is not directly integrated into a judgment but is filtered and weighted by teachers on the basis of, for example, heuristics, control beliefs and beliefs about the subject matter. Beliefs are often rather stable and inflexible, but changes in beliefs can sometimes be initiated through interventions and professional development (Five & Buehls, 2012; Pajares, 1992).

Teacher beliefs about giftedness are essential to their judgments of whether a given student is gifted (Megay-Nespoli, 2001). Sternberg and Zhang (1995; Zhang & Sternberg, 1998) proposed the Pentagonal theory of giftedness comprising five criteria that underlie judgments about giftedness (Sternberg et al., 2011): First, a gifted person has to be excellent in one or more domains (*excellence criterion*). In particular, she or he has to be superior in comparison to peers, not in comparison to, for example, persons who have had considerably more opportunities to learn and practice. Concretizing the excellence criterion, second, the excellent characteristic has to be rare in comparison to other peers (*rarity criterion*). If many students excel in a problem-solving task, for example, none of them would be seen as gifted. Third, the superior ability must (have the potential to) lead to productivity (*productivity criterion*). Persons who only look extremely good are typically not considered gifted, because this trait by itself is not seen as (potentially) productive. Fourth, a person has to demonstrate her or his superiority in a domain via at least one valid test (*demonstrability criterion*). Fifth, the superior ability has to be socially valued to be considered a gift (*value criterion*). A swindler might be very good at what she or he does, but most people will not perceive her or him as gifted. However, other swindlers might see that person as gifted. Support for the theory came from American college students and parents of gifted persons (Sternberg & Zhang, 1995) and from student teachers and school teachers in Hong Kong (Zhang & Sternberg, 1998) who labeled students as gifted or not gifted based on case vignettes.

For the most part, teachers can clearly state which characteristics distinguish gifted students (Hany, 1993). For instance, Schack and Starko (1990) analyzed the preferred indicators of giftedness among preservice, general classroom, and gifted education teachers. All teachers named creativity, the ability to learn quickly and easily, curiosity and learning on one's own initiative as the most influential indicators for their nomination choices. Similar beliefs were reported, for instance, by Endepohls-Ulpe and Ruf (2005), Hany (1997), Lee (1999), Miller (2009), and Persson (1998).

However, teachers could not relate their beliefs about giftedness to scientific conceptions of giftedness in a study by Schroth and Helfer (2009). Around 400 educators, including regular classroom teachers, instructors of gifted children, and (assistant) school principals were asked to state their agreement with different conceptions of giftedness: for example, giftedness as general intellectual ability; as conceptualized by Renzulli's (2005a) three-ring model (i.e., above-average ability, creativity, and task commitment);

or as conceptualized by Sternberg's (1999) Triarchic theory (i.e., analytic, creative, and practical intelligence). All definitions were accepted by the majority of participants. The authors concluded that their participants were confused by the variety of conceptions. However, it could not be clarified whether this confusion was a result of, for instance, poorly-fitting beliefs about giftedness, a lack of knowledge regarding the conceptions mentioned, or an acknowledgement that several conceptions can all be considered valid. To determine teachers' agreement with conceptions about giftedness, a further step can be not to ask about specific conceptions directly but to ask about key characteristics that differ across different conceptions.

Matching the systematization of conceptions of giftedness presented in Section 1.1 (see also Table 1.1, p. 16), research on beliefs about giftedness indicates, first, that teachers typically see giftedness as high intellectual potential to learn and achieve (Endepohls-Ulpe & Ruf, 2005; Moon & Brighton, 2008). However, their beliefs might be strongly tied to actual achievement, too, as teachers include, for example, students' strong work habits into their beliefs about giftedness (Brigthon et al., 2007) and indicated that students needed to demonstrate their high potential to them in Zhang and Sternberg's (1995) study.

Second, teachers have mentioned aspects of general intelligence such as the abilities to draw conclusions and to see patterns and relationships as well as domain-specific abilities like strong verbal abilities as indicators of giftedness (Endepohls-Ulpe & Ruf, 2005; Miller, 2009; Schack & Starko, 1990). Third, they seem to see giftedness as specific to concrete domains, although a holistic view that gifted students are superior in all areas was expressed by one of 16 participants in an interview study by Lee (1999) as well.

Fourth, teachers stated rather multidimensional beliefs about giftedness (Brown et al., 2005; Brighton et al., 2007). They saw characteristics related to intelligence, creativity, and motivation as central for giftedness and typically indicated that gifted students have high values on these characteristics (Endepohls-Ulpe, 2005; Lee, 1999). For example, Miller (2009) analyzed concept maps of elementary school teachers' beliefs about the nature of giftedness. He found that the central characteristics were related to fluid and crystallized intelligence (the abilities to draw conclusions, to see patterns and relationships, an extensive vocabulary and a broad range of knowledge), creative thinking (being imaginative, having original ideas, the ability to find new uses for things), and motivation (enjoyment of experimenting and discovery, boredom when unchallenged).

Fifth, some teachers seem to include environmental variables like parents' educational support and socioeconomic status indirectly into their beliefs about giftedness through the inclusion of characteristics like strong work habits, a broad range of knowledge and an extensive vocabulary (Brighton et al., 2007; Miller, 2009). Scholars of modern conceptions of giftedness use environmental variables in an adaptive manner to indicate, for example, the different work habits or levels of knowledge of gifted students in relation to environmental conditions (Gagné, 2005, Subotnik et al., 2011). In contrast, some teachers' beliefs lead to the inclusion of environmental factors in a rather fixed manner, meaning that students with certain environmental conditions are more likely to be seen as gifted (see McBee, 2006; Peterson & Margolin, 1997).

Sixth, some teachers see giftedness as a stable trait (Reis & Renzulli, 2009; Worrell, 2009), but the belief that giftedness is achievable through deliberate practice has also been mentioned (Ericsson, Nandagopal, & Roring, 2005; Hambrick et al., 2014). Jones, Bryant, Snyder, and Malone (2012) showed that about three-quarters of teachers viewed intelligence as rather changeable instead of fixed. However, although intelligence is a central facet of giftedness for teachers (Endepohls-Ulpe & Ruf, 2005), the relation between their beliefs about the changeability or stability of giftedness and intelligence has rarely been researched. Seventh, teachers indicated that gifted students need to be superior in relation to peers (Zhang & Sternberg, 1998). However, they seem to differ in their beliefs about how many students are gifted. Dahme (1996, as cited by Freeman, 2005) investigated teachers from Germany, Indonesia, and the US and found that German teachers believed that 3.5% of students are gifted. The American teachers estimated 6.4% and the Indonesian teachers indicated that 17.4% of students are gifted. Eighth, research indicates that some teachers see gifted individuals as a rather homogeneous group (Bain, Choate, & Bliss, 2006; Baudson & Preckel, 2013a, 2016). Teachers mentioned quantitative differences like higher cognitive abilities (Endepohls-Ulpe & Ruf, 2005; Lee, 1999), but also qualitative differences like the "spark factor," meaning that gifted students show behavior that can be either positive or negative but is high in intensity, of unusual nature, and visible (Rohrer, 1995).

Although research findings on teacher beliefs about giftedness can be sorted along key characteristics of scientific conceptions about giftedness, there has not yet been a systematic investigation. The present dissertation, therefore, closes this gap.

1.3.3 Teacher variables

Several teacher factors that might influence teachers' judgments and beliefs about giftedness have been researched, and some of them are discussed in the following sections: teachers' experience in general classrooms, experience with gifted students and training in the area of giftedness.

1.3.3.1 Teaching experience in general classrooms

Some differences in teacher judgments about whether or not a student is gifted have been found in relation to their tenure. Siegle et al. (2010) prepared eleven case vignettes with different embedded giftedness characteristics. Student teachers and school teachers were to state whether they would or would not refer the student to a gifted program. They found that school teachers were more likely to nominate a student to the gifted program than student teachers. Similar results were reported by Guskin, Peng, and Simon (1992). However, in both studies, school teachers had experience in the area of giftedness, thereby confounding the results. In Rubenzer and Twaite's (1979) study with 1,220 participants, teachers with six or more years of teaching experience in general classrooms indicated more often than teachers with less experience that they had gifted students in their classes and that the identification of gifted students is not difficult.

However, teaching experience in general classrooms might rarely be relevant for differences in beliefs about giftedness. Specifically, Şahin and Düzen (1994) identified similarities rather than differences between university students in social sciences (e.g., psychology, education, and history) and elementary school teachers' beliefs about giftedness. Similarly, Guskin, Peng, and Majd-Jabbari (1988) reported that student teachers and school teachers agreed on the same categories of giftedness (e.g., verbal ability and creative arts). Moreover, teachers' beliefs about giftedness also seem to be rather stable throughout their careers. Baudson and Preckel (2013a, 2016) did not find differences between student teachers and school teachers or between school teachers with different amounts of teaching experience: Overall, all study participants were in favor of the disharmony hypothesis (i.e., social-emotional maladjustment) instead of the harmony hypothesis (i.e., superiority in every respect) regarding giftedness. The number of years of teaching in general classrooms mostly did not impact teachers' beliefs about gifted education either (Bégin & Gagné, 1994; Chessman, 2010; Cramond & Martin, 1987). An exception was that inexperienced teachers in Rubenzer and Twaite's (1979) study were more in favor of enrichment than experienced teachers.

1.3.3.2 Experience with gifted students

Almost all elementary school teachers in Germany should have—knowingly or unknowingly—experience with gifted students as students of all ability levels are taught in the same classroom at this school level. However, only 19% of the elementary school teachers in Heller, Reimann, et al.'s (2005) study suspected having gifted students in their classes. If teachers know about having gifted students in their school, they seem to be more accurate in judging students' giftedness because they might use their experience with these students as a basis for their judgments about other students' giftedness (Anastasiow, 1964a, 1964b). Based on case vignettes with embedded giftedness characteristics, Bianco and Leech (2010) showed that teachers of gifted classes nominated students as gifted more often than general classroom teachers or special education teachers—independently of whether or not the students had an additional disability label.

Whereas some studies reported no differences in beliefs about giftedness among teachers who stated that they had or did not have experience with gifted students (McCoach & Siegle, 2007; Schroth & Helfer, 2009), more studies indicated differences. Of the 384 elementary school teachers who were asked by Endepohls-Ulpe and Ruf (2005) to describe a gifted student, 50 percent stated that they never had instructed a gifted student. Teachers with experience had more precise and realistic conceptions of giftedness than teachers without experience. Copenhaver and McIntyre (1992) indicated that whether a teacher had taught or not taught gifted classes explained differences in beliefs about giftedness, but not the duration of teaching gifted classes. Teachers without experience in teaching gifted classes, as opposed to teachers with this experience, more heavily stressed that gifted students have a mature personality, are successful in school, exhibit negative characteristics (e.g., lazy, rebellious, inattentive), and need challenges. Also, Schack and Starko (1990) reported differences in relation to having or not having experience in teaching gifted classes. Vocabulary, IQ scores, and multiple interests were more preferred among teachers of gifted classes, whereas grades, class performance, and motivation were more often chosen as indicators of giftedness by student teachers and general classroom teachers. The authors concluded that teachers of gifted classes' beliefs about giftedness were more in line with giftedness indicators proposed by giftedness researchers, whereas the indicators mentioned by preservice teachers and classroom teachers were more in line with school success.

1.3.3.3 Training in the area of giftedness

Teacher training can be associated with teacher judgments of giftedness. Gear (1978) and Şahin and Cetinkaya (2015) developed trainings to, among other things, explain the importance of intelligence and increase teachers' attention to indicators of giftedness. In comparison to teachers who did not participate in these trainings, teachers with training identified more students who had high scores in intelligence tests as gifted. Borland (1978) and Jacobs (1972) also indicated that trained teachers identified gifted children more accurately than untrained teachers. Rubenzer and Twaite (1979) reported that teachers who had received training in the area of giftedness judged that students in their classes were gifted more often than teachers without training. Siegle and Powell (2004) found similar results. Furthermore, they nominated students who had high mental computation and problem solving skills and single areas of interest more often.

The research is rather inconclusive about the effects of training on teachers' beliefs about giftedness: Teachers with training associated fewer negative characteristics with gifted students (Copenhaver & McIntyre, 1992) and reported greater understanding of giftedness and gifted education (Cashion & Sullenger, 2000; Hansen & Feldhusen, 1994) than teachers without training. In a qualitative study, Goodnough (2000) evaluated the change in six student teachers' beliefs about the nature of giftedness over the course of a university course on giftedness that aimed to foster liberal beliefs about giftedness. Three student teachers started with narrow views: for example, that giftedness is a high test score on an intelligence test or that gifted students are geniuses who are superior in many ways. The other three student teachers already had broad multidimensional views of giftedness. All of them expressed broad multidimensional beliefs about the meaning of giftedness at the end of the course. However, having or not having training in the area of giftedness did not always have effects. The teachers in Miller's (2009) study had different levels of gifted education training but did not differ, for example, in the complexity of their beliefs about the nature of giftedness, although a strong heterogeneity of beliefs was observed. In McCoach and Siegle's (2007) study, as well, teachers' beliefs about giftedness did not differ in relation to having training experiences or not.

Copenhaver and McIntyre (1992) reported that teachers who had the most training in the area of giftedness were also those with the most experience in teaching gifted students. This result emphasizes the problem of confounded variables and the possibility of non-representative samples due to interest or personal affectedness. Furthermore, the presented studies did not sufficiently distinguish between different kinds of training or describe the training (e.g., length, goals, content) so as to allow for conclusions about what leads to differences and what is ineffective.

1.3.4 Students' demographic variables

In addition to teacher variables, the influence of several student background variables on teacher judgments and beliefs about giftedness has been investigated. The following section addresses students' age, which is confounded with the level of school teachers teach (i.e., elementary vs. secondary school level), as well as students' gender and social class.

1.3.4.1 Age: elementary vs. secondary school level

Fatouros (1986) and Jacobs (1972) argued that younger children, such as preschool children, might be more difficult to identify than older children. However, in Acar et al.'s (2016) meta-analysis, the accuracy of teacher judgments of giftedness was not significantly different among preschool, elementary school, or secondary school teachers. Furthermore, the (non-significant) tendency was for higher accuracy levels among younger students than older ones. Acar and colleagues reasoned that elementary school teachers have more occasions to observe students than secondary school teachers.

Beliefs about giftedness among elementary and secondary school teachers seem to be rather similar in many aspects, although differences have also been observed: Elementary and secondary school teachers both connect high cognitive abilities and the willingness to learn with giftedness (Endepohls-Ulpe, 2005; Hany, 1997; Lee, 1999; Persson, 1998). Creativity was mentioned by elementary school teachers (Miller, 2009), but was not a central giftedness characteristic for secondary school teachers (Hany, 1997). In a study by Copenhaver and McIntyre (1992), both elementary and secondary school teachers emphasized the inquisitive minds of gifted students. However, elementary school level teachers identified negative characteristics and an extensive vocabulary as typical for gifted students more often than secondary school teachers. The authors reasoned that elementary school teachers have more heterogeneous groups and, therefore, might notice these characteristics more easily. However, concerning negative characteristics, Baudson and Preckel (2013a, 2016) showed that elementary and secondary school teachers both supported a disharmony hypothesis of giftedness, meaning that they believed gifted students were intellectually superior but with emotional and social maladjustments. However, as students' age and the level of school teachers teach are confounded, studies about whether teachers at different school levels see differences in giftedness across the lifespan are rare (e.g., elementary school teachers' beliefs about giftedness in secondary school students or secondary school teachers' beliefs about elementary school students). Therefore, more research is needed whether teachers' beliefs differ in relation to students' age.

1.3.4.2 Gender

Teacher judgments of whether or not a student is gifted often tend to favor boys over girls, although girls seem to obtain better school grades (Duckworth & Seligman, 2006). For example, seven percent more boys were nominated as gifted in a study by Hernández-Torrano, Prieto, Ferrándiz, Bermejo, and Sáinz (2013). When teachers were asked to describe gifted students (Endepohls-Ulpe & Ruf, 2005) or indicate whether they have taught gifted students (Heller, Reimann, et al., 2005), teachers named a boy twice as often than a girl. Moreover, when teachers were asked to nominate fictitious students for gifted education based on vignettes, Bianco, Harris, Garrison-Wade, and Leech (2011) reported that more boys than girls were nominated as gifted. However, in Siegle and Powell's (2004) study with case vignettes, students' gender was not associated with teachers' nomination decisions for gifted education.

Students' gender seems to be connected to teachers' beliefs about giftedness. In Baudson and Preckel's (2016) study, which used case vignettes as well, gifted and average-ability boys were rated as more intelligent than gifted and average-ability girls. As the authors argued, these gender differences might be in part related to findings that teachers tend to attribute boys' success more to (innate) ability and girls' school achievements more to effort (Fennema, Peterson, Carpenter, & Lubinski, 1990; Siegle & Reis, 1998). Bianco et al. (2011) used identical vignettes that differed only with regard to students' gender and found that teachers perceived the girl's social competence as lower than the boy's. Teachers thus believed that the girl was not ready for gifted education. However, Endepohls-Ulpe (2008) did not find gender differences in teachers' ratings of vignette cases. Instead, teachers associated social maladaptation with giftedness. But the author stated that gender stereotypes impact students' social behavior and thereby indirectly affected teachers' beliefs about giftedness.

Results are inconsistent concerning beliefs about whether boys and girls are differently gifted in different domains: For example, girls were believed to have higher verbal abilities than boys (Hinnant, O'Brien, & Ghazarian, 2009; Siegle & Reis, 1998). In return, boys were believed to have higher ability in mathematics than girls, although this result was found in only some studies (e.g., Li, 1999; Ziegler, Kuhn, & Heller, 1998), but not in others (e.g., Hinnant et al., 2009; Siegle & Reis, 1998). Furthermore, Siegle (2001) indicated that teachers tend to see giftedness more in students who do not match expected gender stereotypes.

1.3.4.3 Social (i.e., socioeconomic and ethnic) background

Teacher judgments about giftedness seem to differ in relation to students' socioeconomic background, thereby disadvantaging students from families with low socioeconomic status: McBee (2006) had a dataset containing data for all public school students in a US state during the year 2004. In his study, the accuracy of teacher judgments was lower for low SES students (i.e., who received free or price-reduced lunch) than for high SES students (i.e., whose parents paid for lunch). Brighton et al. (2007) stated that the following beliefs and foci are reasons why gifted students from low socioeconomic backgrounds and with less preschool educational experience are often overlooked by teachers: Teachers often hold rather traditional beliefs about giftedness that include a broad range of knowledge, an extensive vocabulary, and the ability to work independently (Miller, 2009; Moon & Brighton, 2008; Speirs Neumeister, Adams, Pierce, Cassady, & Dixon, 2007). They direct their attention more strongly towards students' deficits than their strengths (Brighton, Hertberg, Moon, Tomlinson, & Callahan, 2005). Furthermore, their expectations about students' academic achievement are connected to students' behavior, dress, speech patterns, and the congruence of parents' communicated education-related values with the values of the teachers (Brighton et al., 2007).

With respect to different cultural backgrounds, the aforementioned study by McBee (2006) indicated that US teachers were more accurate in nominating Asian American, Native American, and Caucasian students as gifted than Hispanic and African American students, which might be due to racism or cultural ignorance. Specifically, cultural ignorance was debated in the context of teachers' beliefs about giftedness: Teachers rarely consider environmental and cultural factors in their views of giftedness and might therefore not be sensitive to expressions of giftedness among students from certain cultures that are not their own (Speirs Neumeister et al., 2007). Peterson and Margolin (1997) had middle school teachers who were mainly Caucasian discuss which students they would nominate for a gifted education program. Comments about students' verbal skills, personality and family status stood out. Many students from minority groups were overlooked because teachers used the ideals and values of the dominant Caucasian culture in the US to guide their judgments of giftedness. They generalized the values of the mainstream group across all groups without awareness that the ideals, values, and thereby the expressions of giftedness among some minorities might be different.

1.3.5 Facets of students' giftedness

Giftedness is mostly seen as a multifaceted construct and significant facets of giftedness are students' (potential for) high academic achievement, intelligence, creativity, and motivation (see Sternberg & Davidson, 2005 and Section 1.1). Elementary school teachers also reported that they consider these characteristics as relevant for giftedness (e.g. Miller, 2009, see Section 1.3.2). Neber (2004) investigated which student characteristics influenced teacher judgments about giftedness. He asked teachers to nominate all highly able students who qualified for a gifted education program. These students were then tested and questioned concerning several characteristics. In addition, teachers rated students' characteristics and indicated which of the students were most qualified to participate in the program. The students who were rated as most qualified had, in comparison to students who were rated as highly able, higher test scores for cognitive abilities and were rated higher by their teachers in their motivation, learning skills, communication skills, and school achievement. Based on a regression analysis, teachers' decisions about the most qualified students were predicted by test scores for cognitive abilities, teacher ratings of communication skills, and teacher ratings of interest in school subjects, but not by school grades. However, the authors did not mention the correlations between school grades, teacher ratings, and test scores, so concern about multicollinearity remains. Hany (1997) argued that teachers are trained to judge students' achievements, not their potential. Specifically, a study by Rost and Hanses (1997) showed that teachers primarily tend to identify gifted students as gifted if they have good school grades, thereby overlooking underachievers.

Teachers' accuracy in judging facets of giftedness needs to be viewed in light of the importance of teacher ratings of student characteristics for their judgments about giftedness. Urhahne (2011) investigated whether teachers can identify students in relation to Renzulli's (2005a) three-ring conception of giftedness. To do so, he compared teacher ratings with students' test scores and self-reports concerning students' cognitive abilities, creativity, and motivation. Teachers' effectiveness was 50% and their efficiency was 25.6% when it came to identifying students with high levels in all three characteristics. The author concluded that teachers might not be suited for the identification of gifted students when a multidimensional giftedness conception is the basis. Overall, there does not seem to be a general diagnostic competence to accurately rate students' characteristics (Spinath, 2005). Teachers have been found to be more accurate in rating students' academic achievement and intelligence and less accurate in rating students' creativity and motivation (Schrader, 2010; Sommer, Fink, & Neubauer, 2008; Spinath, 2005; Südkamp et al., 2012; Urhahne, 2011). However, not only ratings of academic achievement but also all other ratings seem to be connected to students' academic achievement (e.g., Kaiser, Retelsdorf, Südkamp, & Möller, 2013; Gralewski & Karwowski, 2013; Machts et al., 2016). Furthermore, the correlations between ratings of facets of giftedness were found to be higher than the correlations between student data such as different tests and self-reports (e.g., Sommer et al., 2008; Urhahne, 2011). This effect has been discussed under the banner of *halo effects* (e.g., Babad, Inbar, & Rosenthal, 1982; Urhahne, 2015). The following section is a short overview of the accuracy levels of teacher judgments about academic achievement, intelligence, creativity, and motivation. Research on halo effects is also presented.

1.3.5.1 Academic achievement

Hoge and Coladarci (1989) analyzed 16 studies that investigated teacher judgments of academic achievement. The associations between judgments and student achievement varied greatly, from weak to strong, with a median correlation of $r = .66$. Studies in which teachers made direct judgments about students' achievement in a test reported higher correlations than studies in which the content of the rating did not match the content of the test. Supporting results were presented in a more recent meta-analysis by Südkamp et al. (2012). They looked at 75 studies that focused on teachers' judgments of students' academic achievement and that were published in the last 20 years. In general, the association was relatively high, with a mean correlation of r = .63. The effect sizes fluctuated greatly between studies. Two significant moderators were found. First, teachers' judgment accuracy was lower when teachers were not informed about the criterion against which their judgments were compared. Second, the accuracy of judgments was lower when the content of the judgment and the test differed.

1.3.5.2 Intelligence

A recent meta-analysis (Machts et al., 2016) reported a strong correlation between ratings and test scores for intelligence (r = .50) across 19 studies. The association was higher when teachers received a frame of reference (e.g., instructions to compare students to same-age students or to rate whether students would solve concrete items on an intelligence test correctly) and when the multi-level structure of teacher ratings was accounted for (i.e., that teacher ratings of students in the same class are likely to be more similar to each other than ratings of students from different classes). However, the authors also found that teacher ratings of intelligence were more highly correlated with students' academic achievement than with their intelligence test scores. Fischbach et al. (2013, p. 116), whose study was included in this meta-analysis, concluded that "teachers were hardly able to accurately judge their students' intelligence above and beyond the part of intelligence that was shared by IQ scores and GPA [General Point Average]". However, the correlation between intelligence tests and school grades is indeed high, with ρ = .54, but far from perfect (Roth et al., 2015). Miller and Davis (1992) indicated that the accuracy of teacher judgments is higher for students' verbal, mathematical and nonverbal abilities than for students' memory and that judgment accuracy is not dependent on students' level of cognitive ability. Focusing on ratings of high ability, the teachers in Heller, Reimann, et al.'s (2005) study had to indicate whether or not a student belonged to the 10% most intelligent students in a class. These ratings correlated more highly with students' test scores in verbal and mathematical reasoning and the total score for intelligence than with test scores for nonverbal reasoning.

1.3.5.3 Creativity

In relation to teacher ratings of students' creativity, the meta-analysis by Machts et al. (2016) that included six studies about creativity reported a mediocre association of $r = .36$ for teacher ratings and test scores. The distribution of the reported strengths of association, however, was rather skewed: four of the six studies reported low associations. Two of the studies that reported low accuracy levels used only one item as a teacher rating of students' creativity (Gralewski & Karwowski, 2013; Urhahne, 2011). Concerning the remaining two studies, Sommer et al. (2008) reported a mediocre and García-Ros, Talaya, and Pérez-González (2012) a strong correlation. Both studies developed scales with items about creativity that were seen as relevant for high ability and giftedness and were in alignment with the tested content. Another study that was not included in the meta-analysis investigated whether teachers can identify students with high levels of creativity (Heller, Reimann, et al., 2005). They asked with one item whether students belonged to the most creative 10% of students in the class and found low associations with test scores of creativity. It was argued that the strong variability in findings might be in part attributed to the presence or lack of information about the frame of reference and the content of creativity (Machts et al., 2016). Furthermore, creativity judgments seem to be dependent on students' school performance. Gralewski and Karwowski (2013) reported that teachers tended to judge students who were good at school as creative.

1.3.5.4 Motivation

Teachers' accuracy in rating students' motivational characteristics seems to be low to moderate (Dicke, Lüdtke, Trautwein, Nagy, & Nagy, 2012; Givvin, Angeles, Stipek, Salmon & MacGyvers, 2001; Harvey, Suizzo, & Jackson, 2016; Spinath, 2005). For instance, Gagné and St Père (2001) reported correlations between $r = .03$ and $r = .24$ with students' self-reports of intrinsic and extrinsic motivation and persistence. Skinner, Kindermann, and Furrer (2009) found—with correlations ranging from $r = .21$ to $r = .44$—small to medium connections between teacher ratings and students' self-reports regarding different aspects of engagement and disaffection. The correlations were higher for behavioral indicators of engagement and disaffection than for not as easily observable indicators and for aggregated scales of engagement and disaffection. Kaiser et al. (2013) also indicated that teacher judgments of students' engagement were connected to students' achievement. The finding of their field study and two simulation studies showed that students' engagement and academic achievement were both predictive of teachers' judgments of engagement.

1.3.5.5 Halo effects

When teachers have to rate several facets of giftedness for the same student, the correlations between rated characteristics are often higher than between corresponding student characteristics as measured by tests or self-reports (Chan, 2000; Pfeiffer, Petscher, & Kumtepe, 2008; Sommer et al., 2008; Wild, 1993). For example, Urhahne (2011) showed that teacher ratings of students' mathematical ability, creativity, and task commitment seem to be influenced by halo effects. Whereas student characteristics correlated between r = -.07 (mathematical abilities and task commitment) and r = .25 (mathematical abilities and creativity), teacher ratings ranged between r = .54 (mathematical abilities and creativity) and r = .67 (mathematical ability and task commitment). This phenomenon is discussed under the banner of *halo effects* (e.g., Babad, Bernieri, & Rosenthal, 1989; Babad et al., 1982; Burke, Haworth, & Ware, 1982; Mason, Gunersel, & Ney, 2014; Urhahne, 2015). Thorndike (1920) already noticed that "… the estimates of the same man in a number of different traits […] were very highly correlated and very evenly correlated" (p. 25). He named that effect "halo" and argued that this constant error might be due to the influence of a general impression of a person on judgments regarding different characteristics of that person or due to one salient characteristic that affects the judgment of other characteristics. Fisicaro and Lance (1990) added one further possibility: that raters might be unable to conceptually discriminate among the different dimensions.

Debated reasons for halo effects on facets of giftedness often concern teachers' strong reliance on students' school achievement: Neber (2004) stated that students' school achievement is often generalized too strongly to other student characteristics. Applying the *GRS* (*Gifted Rating Scales*; Pfeiffer & Jarosewich, 2003), Li et al. (2009) reported high correlations between ratings of, for example, intellectual ability, academic ability, creativity, and motivation, but they could still establish a factor structure with separable factors for the rated characteristics. However, Anders, McElvany, and Baumert (2010) found a global factor for academic achievement that included ratings of cognitive abilities and motivation. Similarly, Burke et al. (1982) reported a factor that included ratings of learning and academic abilities. The authors discussed that this factor might be decisive in judgments of giftedness rather than other factors, which related to ratings of, for example, creativity and leadership.

1.3.6 Student characteristics on the class or school level: Reference group effects

Specifically in the area of giftedness, the consideration of reference groups is a relevant topic. The construct of giftedness itself is dependent on a reference group and, furthermore, whether to compare students with peers, highly able individuals in a domain, national norms, or local norms are frequently discussed issues (e.g., Lohman & Gambrell, 2012; Subotnik et al., 2011, 2012; see Section 1.1). In addition to the consideration of reference groups on a theoretical level for the construct of giftedness, attention needs to be directed to how teachers' beliefs and judgments about giftedness are connected to reference groups.

Relatively little research is available investigating how teachers' views of who is gifted are influenced by reference groups. Lee (1999) and Zhang and Sternberg (1995) noted that teachers see gifted students as excellent and with rare abilities in one or more domains in comparison to other peers. Anastasiow (1964a; 1964b) had already investigated frames of reference for teacher nominations of gifted students. He argued that teachers of high achieving classes might be more accurate in their judgments of giftedness than teachers of classes with lower average levels of achievement. He reasoned that teachers might orient themselves on the class mean of achievement and nominate students above that mean. The probability is higher that there are actually more gifted students in classes with high average achievement levels (i.e., higher base rate) than in classes with lower mean achievement. Hence, teachers' chance of correctly identifying a gifted student increases in classes with high average achievement levels. However, based on his empirical study of teachers who were to nominate students as gifted (Anastasiow, 1964b), he concluded that teachers' accuracy is more dependent on the number of previously identified gifted students in a school than on the achievement level in a class. McBee (2010) also investigated composition effects on the identification of elementary school students as gifted. Identification was based on a multi-criterial assessment including intelligence, creativity, and motivation. To be assessed for giftedness, students had to be nominated as potentially gifted. One of the nomination possibilities was teacher nomination. He found that the composition of students in a school was associated with the identification rate. Similar to Anastasiow's (1964b) findings and reasoning, students had a higher probability of being identified when they came from schools with a high percentage of students previously identified as gifted and that had a highly academic environment (i.e., high percentage of students scoring "advanced" on an academic achievement test). Furthermore, the chances of getting identified as gifted were higher in schools with high percentages of Asian American students, with low percentages of negative behavior incidences per student (e.g., drug and weapons possession), and with low percentages of students who have been retained. However, the composition effects were not separately tested for teacher nominations and—in the absence of suitable variables—the author did not control for students' individual levels of achievement or intelligence.

Reference group effects have been investigated in other research areas as well. For example, gifted students' academic self-concept differs in relation to class-average or school-average abilities (see Section 1.2.2). But importantly, the negative reference group effect has also been found to affect teacher ratings: In several studies, a student received lower teacher-assigned grades if she or he was in a class with a higher average achievement level than a student with comparable abilities in a class with a lower average achievement level (Marsh, 1987; Trautwein & Baeriswyl, 2007; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). These results were supported by an experimental study by Südkamp and Möller (2009), but they found an effect only for school grades, not for an estimation of correct answers. In a study by Piopiunik and Schlotter (2012), this "grading on a curve" was only detected for female teachers, not male teachers, and was independent of the level of student performance, teacher experience, and class size.

Negative reference group effects have also been found in the tracked school systems in Germany and Switzerland, in which elementary school teachers typically give a recommendation (i.e., transition recommendation) about which of the secondary school tracks separated by achievement level students should attend (Baeriswyl, Wandeler, & Trautwein, 2011; Tiedemann & Billmann-Mahecha, 2007; Wagner, Helmke, & Schrader, 2009). Students received transition recommendations for higher-ability school tracks if they belonged to classes with lower levels of average achievement than students with the same achievement level who belonged to classes with higher levels of average achievement. Milek, Lüdtke, Trautwein, Maaz, and Stubbe (2009) showed that teachers in all German states that were examined in this study were influenced by a negative reference group effect and that there were no statistically significant differences between

German states. However, the negative reference group effect on transition recommendations seemed to be mediated by school grades (Gröhlich & Guill, 2009; Trautwein & Baeriswyl, 2007).

Furthermore, Baudson et al. (2014) focused on teacher judgments of elementary school students' cognitive abilities (i.e., first to third grade). Again, reference group effects were observed: students from classes with high class intelligence means were judged lower than students with the same intelligence scores but from classes with lower class mean levels of intelligence. Similarly, Trautwein and Baeriswyl (2007) found that teachers rated the cognitive abilities of students with equal achievement more negatively in classes with higher class mean achievement levels. Teacher ratings of students' motivation were not influenced by the classes' average achievement level.

Overall, students' academic abilities are important for teacher ratings of these characteristics both individually and on the class level. Students' academic abilities are significant ingredients of giftedness definitions and teachers' beliefs about giftedness. Therefore, it seems plausible to expect a negative reference group effect on teacher judgments about who is or is not gifted. This hypothesis will be addressed in this dissertation.

1.3.7 Teacher judgments of (facets of) giftedness in comparison with parent judgments

Teachers are not the only significant adult agents for (gifted) students' academic development. Students' parents have to be accounted for as well. In particular, how teachers and parents concur in learning-related issues like values but also in their expectations and ratings of learning-relevant characteristics is considered to be crucial for students' academic development (Brenner & Mistry, 2007; Glueck & Reschly, 2014; Peet et al., 1997). Therefore, their judgments of giftedness and facets of giftedness are compared in this section.

Teachers' and parents' perspectives on students are different because in contrast to parents, teachers have typically more professional knowledge about education and academic development, see children in a larger number and variety of academically relevant situations and can compare them with a larger group of other children (Harder, Trottler, Vialle, & Ziegler, 2015; Petscher & Li, 2008). Teachers have higher accuracy levels in identifying gifted students (McBee, 2006), particularly if they use rating scales

(Acar et al., 2016). However, Gross (1999) and Jacobs (1971) reasoned that parents are better at identifying giftedness in kindergarten children and students in the first few years of schooling.

Whether teachers are more, similar, or less accurate in rating different facets of giftedness seems to vary depending on the rated characteristics: Based on Brenner and Mistry's (2007) study, teachers might be more accurate than parents in judging students' academic achievements. Teachers' accuracy in rating students' academic achievement is high on average (Südkamp et al., 2012), and Schrader (2010) summarized that the accuracy of parent ratings of academic achievement is low to high. Furthermore, teachers seem to be similar or more accurate in judging students' cognitive abilities: Teachers' and parents' accuracy in rating intelligence seems to be high (e.g., Machts et al., 2016; Rennen-Allhoff, 1991) and was similarly high in a study by Sommer et al. (2008). However, a higher convergent validity for teacher ratings of analytical abilities was found by Geiser, Mandelman, Tan, and Grigorenko (2016). Miller and Davis (1992) reported higher correlations between ratings and test scores for teachers than parents with regard to verbal, mathematical, and figural abilities but concluded based on the comparison of difference scores (i.e., differences between ratings and test scores) that teachers and parents were equally accurate in their judgments. With regard to ratings of students' creativity, teachers were found to be more accurate than parents (Geiser et al., 2016). Whereas teacher ratings' accuracy was mediocre on average (Machts et al., 2016), low accuracy levels have been reported for parents (Sommer et al., 2008). Focusing on ratings of school-related motivation, teachers might be less or similarly accurate compared to parents, although studies that directly compare both ratings are rare. Low to medium correlations of ratings with self-reports have been found for teachers (e.g., Dicke et al., 2012; Helmke & Schrader, 1989; Spinath, 2005) and low to high correlations for parents (Genser, Straser, & Garbe, 1981; Helmke & Schrader, 1989).

Parent ratings might be affected by halo effects just as teacher ratings are, but to a lesser degree: For example, Sommer et al. (2008) reported that teacher and parent ratings of intelligence and creativity were highly correlated, whereas the corresponding test scores had a much lower correlation. Other studies also reported high correlations between ratings of different characteristics for teachers and parents, but stated on a descriptive level that the correlations between teacher ratings were higher than between parent ratings (Petscher & Li, 2008; Pfeiffer et al., 2008). Chan (2000), for example, indicated that teacher ratings of students' learning skills, ability in mathematics and science, creativity, leadership skills, and motivation were more highly correlated ($r = .39$ to $r = .71$) than parent ratings ($r = .24$ to $r = .65$). However, Chan (2000), Petscher and Li (2008), and Pfeiffer et al. (2008) did not include student data or did not specify how strongly student variables were correlated with each other.

The congruence between teachers and parents seems to vary strongly depending on the rated characteristic: Teacher and parent ratings of academic achievement were moderately correlated in Brenner and Mistry's (2007) study. Medium to high correlations have been reported for ratings of students' intelligence (Geiser et al., 2016; Miller & Davis, 1992; Sommer et al., 2008; Spinath & Spinath, 2005). Ratings of motivation seem to be moderately correlated (Chan, 2000; Peet et al., 1997), and weak correlations were reported for ratings of creativity (Geiser et al., 2016; Runco, 1989; Sommer et al., 2008). As the results mentioned above did not focus on teacher-nominated gifted elementary school students, the present dissertation addresses ratings of this target group.

1.4 Research Questions of the Present Dissertation

The present dissertation aims to gain insights into teacher judgments and beliefs about giftedness. Scholars stress the importance of identifying gifted students as early as possible to be able to support their socio-emotional and academic development (Fatouros, 1986; Heller, 2004; Karnes & Johnson, 1990; Schofield & Hotulainen, 2004) and have, furthermore, outlined several areas in which teacher judgments about ability and giftedness are connected to students' development (see Section 1.2). This dissertation therefore specially focuses on elementary school teachers.

The body of research on teacher judgments about giftedness has accumulated a broad range of results concerning the influence of teacher characteristics, student characteristics, judgment characteristics, giftedness criterion characteristics, and different kinds of accuracy measurements (see Section 1.3). Specifically, teachers' beliefs about giftedness have received strong attention as beliefs can filter and frame information and interpretations (Five & Buehl, 2012; Kleber, 1992; Shavelson & Stern, 1972) and might thereby influence teacher judgments about students' giftedness. Still, important questions need to be solved concerning beliefs about giftedness, judgments about students' giftedness as well as judgments about facets of students' giftedness. This dissertation focused particularly on the following open questions: How do teachers position themselves with regard to key aspects of scientific conceptions about giftedness? Are teacher judgments about giftedness connected to negative reference group effects? In comparison to parent judgments, how do teachers judge facets of giftedness among students who they nominate as gifted? Furthermore, how is the congruence between teacher and parent ratings related to students' school achievement?

To tackle these questions, three empirical studies were conducted (see Table 1.2, p. 50). Teachers' beliefs about the nature of giftedness were investigated in Study 1 (*Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness*). The diversity of different definitions of giftedness (see Sternberg & Davidson, 2005) and the frequent debates about key characteristics of giftedness indicate that there is no one definition that all experts in the area of giftedness can agree upon (see Section 1.1). In addition, teachers have different understandings of giftedness (see Section 1.3.2); however, a systematic analysis of teachers' beliefs and how they align with often-discussed aspects of giftedness is still lacking. Therefore, this study tried to close this gap and explored teachers' beliefs in light of dimensions of giftedness that are often discussed in the scientific field. To assess these beliefs, a questionnaire was developed. As a framework to extract relevant dimensions of beliefs about the nature of giftedness, Subotnik et al.'s (2011, 2012) comprehensive conception of giftedness, which is based on a review of conceptions of giftedness, was used. To set elementary school teachers' beliefs about the nature of giftedness into relation with the beliefs of other teachers, 131 elementary school teachers were compared with 529 student teachers and two groups of enrichment program instructors—one group who were also school teachers (N = 212) and one group who were not school teachers but rather experts in the area that they taught in the program (N = 363). The foci in this study were the structure of these beliefs and structural and mean-level differences in these beliefs between the teacher groups. Furthermore, the relations of beliefs about the nature of giftedness with teachers' number of years in a general classroom and in the enrichment program, and beliefs about whether intelligence is fixed or malleable, were investigated.

Elementary school teachers' judgments about students' giftedness were addressed in Study 2 (*Exploring Reference Group Effects on Teachers' Nominations of Gifted Students*) by examining teacher nominations of students for a statewide enrichment program for gifted elementary school students. Previous research has shown that, among several student characteristics, students' intelligence is connected to teacher judgments of whether these students are gifted (Acar et al., 2016; Machts et al., 2016; Neber, 2004; see Section 1.3.4-5). However, only a few studies have considered that student variables that are aggregated on a class level might be relevant for teacher judgments about giftedness (Anastasiow, 1964b; McBee, 2010; see Section 1.3.6). Research on, for example, teacher judgments of students' cognitive abilities has shown that students' cognitive abilities were rated higher by teachers if they were in classes with lower levels of cognitive abilities than students with the same cognitive abilities in classes with higher average levels of cognitive abilities (Baudson et al., 2014). In this study, it was therefore hypothesized that students' individual fluid and crystallized intelligence were positively connected to students' probability of getting nominated. However, after controlling for students' individual intelligence, it was hypothesized that students from classes with higher average levels of fluid or crystallized intelligence would have a lower probability of getting nominated than students from classes with lower average levels of fluid or crystallized intelligence. Furthermore, this study explored whether three teacher variables, namely experience with giftedness, beliefs about the malleability of intelligence, and the belief that giftedness is domain-specific or holistic, were connected to students' probability of getting nominated or with the expected reference group effect. For this purpose, the nomination decisions of 105 elementary school teachers for their 1,468 third graders were viewed.

Investigating elementary school teachers' judgments about facets of giftedness, namely students' verbal abilities, mathematical abilities, deductive reasoning, creative thinking, and engagement, was the content of Study 3 (*A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students*). Specifically, elementary school teachers' ratings of students who they nominated for an enrichment program were compared with parent ratings because, first, research on teacher and parent ratings of teacher-nominated elementary school students is rare (see Section 1.3.7) and, second, the congruence between teacher and parent ratings of students' competence and school engagement has been found to be related to students' academic achievement (Peet et al., 1997, see Section 1.2.3). In this study, the topics of the comparison were (1a) the accuracy of teacher and parent ratings, (1b) whether they were differently affected by halo effects, and (1c) how teacher and parent ratings of the same student characteristics were correlated with one another. Whether the reported effect of teacher-parent congruence in ratings on students' school achievement can also be found for ratings of facets of giftedness was further investigated. Based on previous research (Baudson & Preckel, 2013b; Geiser et al., 2016; Li et al., 2008; Miller & Davis, 1992; Sommer et al., 2008; Petscher & Li, 2008; Pfeiffer et al., 2008) stemming mainly from general classroom samples, the following was expected: (1a) The accuracy of teacher and parent ratings of cognitive abilities is either similar or higher for teachers, is higher for teacher ratings of creative thinking, and does not differ between teachers and parents or is lower for teachers for ratings of engagement. (1b) Both ratings, but teacher ratings more strongly than parent ratings, were affected by halo effects. (1c) Teacher and parent ratings correlated moderately or highly for cognitive abilities, weakly for creative thinking, and moderately for engagement. Furthermore, following Brenner and Mistry (2007), two hypotheses were investigated: (2a) that school grades should be best when teacher and parent ratings were congruently high and worst when teacher and parent ratings were congruently low, and (2b) that high parent ratings would reduce the associations between teacher ratings and school grades.

Although each study had different specific foci, the three studies shared four aspects in particular: first, all studies included—although not exclusively—elementary school teachers from general classrooms. In Germany, students are not segregated by ability but are taught together in one classroom at the elementary school level. Therefore, for all elementary school teachers, it is a possibility that they have taught or will teach gifted students. Second, all samples stemmed from the state of Baden-Württemberg, Germany, in the period from 2012 to 2014, so that, in addition to the school system, many contextual variables like the socio-political setting were similar for all three studies. Third, all studies were imbedded in the context of a statewide enrichment program for gifted elementary school students. The first study included teachers from the enrichment program and general classroom teachers who were aware of this program. In the second study, nominations for the enrichment program were viewed based on the same sample of general classroom teachers from Study 1. In Study 3, teacher and parent ratings of students who participated in the enrichment program were viewed. Fourth, none of the studies transmitted to the participants what giftedness “truly” is, which students are “truly” gifted, or how gifted students “truly” are. The exclusion of these standards was rooted, first, in the goal of this dissertation to capture the beliefs and judgments as unfiltered as possible and, second, in the diversity of definitions of giftedness, which lead to different giftedness criteria (see Sternberg & Davidson, 2005).

In the following three chapters (2-4), the three empirical studies outlined above will be presented in great detail. Chapter 2 will deal with Study 1 *Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness,* Chapter 3 with Study 2 *Exploring Reference Group Effects on Teachers' Nominations of Gifted Students*, and Chapter 4 with Study 3 *A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students.*

Table 1.2

Overview of the Goals, Research Questions, and Samples of the Three Empirical Studies of the Dissertation

Study	Study goal	Research questions	Sample
Study 1: Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness	Exploring teachers' beliefs about the nature of giftedness in alignment with key characteristics of a modern conception of giftedness	1. How are beliefs about the nature of giftedness structured? 2. Will the (a) structure and (b) mean level of beliefs about the nature of giftedness differ between the different groups of teachers? 3. How are beliefs about the nature of giftedness related to beliefs about the malleability of intelligence?	*N* = 131 elementary school teachers *N* = 212 school teachers who worked in an enrichment program *N* = 363 instructors who worked in an enrichment program but were not school teachers *N* = 529 student teachers
Study 2: Exploring Reference Group Effects on Teachers' Nominations of Gifted Students	Exploring the reference group effect of school class-average levels of intelligence on students' probability of getting nominated as gifted	1. Are students' individual levels of fluid and crystallized intelligence positively associated with their probability of being nominated? 2. Are teachers' nominations negatively affected by the class-average levels of fluid and crystallized intelligence, controlling for individual levels of fluid and crystallized intelligence? 3. Are (a) teachers' experience with giftedness, (b) their beliefs about the changeability of intelligence, and (c) their belief that giftedness is holistic or domain-specific associated with the size of the reference group effect or with students' probability of being nominated?	*N* = 105 elementary school teachers *N* = 1,468 third graders
Study 3: A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students	Comparing teacher and parent ratings of facets of giftedness and examining whether teacher-parent congruence is connected to students' school grades	1a. Is the accuracy of teacher ratings in comparison to parent ratings lower or not different for engagement, not different or higher for cognitive abilities, and higher for teacher ratings of creative thinking? 1b. Are both ratings – but teacher ratings more strongly than parent ratings - affected by halo effects? 1c. Are the two ratings correlated weakly for creative thinking, moderately for engagement and moderately or highly for cognitive abilities? 2a. Are school grades best when teacher and parent ratings are congruently high and worst when teacher and parent ratings are congruently low? 2b. Do high parent ratings reduce the associations between teacher ratings and school grades?	*N* = 294 elementary school teachers *N* = 408 third and fourth graders who participated in an enrichment program *N* = 535 parents

2

Study 1:
Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness

Rothenbusch, S., Zettler, I., Lösch, T, & Voss, T. (2016). Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness. *Manuscript submitted for publication.*

Abstract

In this study, we tested to what extent teachers' beliefs about the nature of giftedness are aligned with often discussed characteristics of conceptions of giftedness. Specifically, the focus was on the structure of these beliefs, structural and mean-level differences between these beliefs of different teacher groups, and the relation between these beliefs and beliefs about the malleability of intelligence. A scientific conception of giftedness was used to derive eight dimensions of beliefs about the content and development of giftedness, which empirically supported in the total sample (*N* = 1,235). Partial strong measurement invariance was achieved for comparing 529 student teachers, 131 elementary school teachers, and two groups of teachers from an enrichment program for gifted elementary school students (212 school teachers, 363 instructors). Student teachers' beliefs differed the most from the other groups. Teachers' beliefs about the nature of giftedness were related to their beliefs about the malleability of intelligence.

Keywords: beliefs about giftedness, factor analysis, beliefs about intelligence, measurement invariance testing, teachers' beliefs

Elementary School Teachers', Enrichment Program Teachers', and Student Teachers' Beliefs About the Nature of Giftedness

2.1 Introduction

Researchers have been interested in the construct of giftedness for decades (Stoeger, 2009). As it is the case for many other constructs, however, the research community has not reached a consensus on what giftedness is (Carman, 2013). Freeman (2005), for instance, estimated that approximately 100 definitions of giftedness exist. Although giftedness is typically viewed in relation to reference groups such as high-functioning individuals in a domain (Subotnik, Olszewski-Kubilius, & Worrell, 2011, 2012), conceptions differ, for example, in their emphasis on intelligence and achievement, their considerations of developmental issues, or their inclusion of additional personal and environmental factors (see Shavinina, 2009). These and other differences in definitions can result in the identification of different students as gifted or in different approaches for providing educational support for gifted students (see Sternberg & Davidson, 2005).

When already researchers do not agree on one definition, school teachers and instructors of gifted students might have different understandings of giftedness, too. Teachers play a significant role in both the identification and development of gifted children. For instance, teachers are often involved in the process of identifying gifted students (e.g., McClain & Pfeiffer, 2012; Miller, 2005; Rothenbusch, Zettler, Voss, Lösch, & Trautwein, 2016; Schack & Starko, 1990). Also, virtually all teachers of regular school classes will teach gifted students throughout their careers because gifted students can be found in all age groups, classes, and regions (Bangel, Moon, & Capobianco, 2010; Ross, 1993). Knowing teachers' beliefs about the nature of giftedness might help researchers understand teachers' perceptions of and behaviors toward gifted students as well as other educational-related beliefs that teachers might hold about gifted students. Furthermore, if these beliefs are known, training programs in the area of giftedness can be more specifically tailored to the target group.

In the present study, we analyzed beliefs about the nature of giftedness held by elementary school teachers, two groups of teachers working in an enrichment program for gifted elementary school students, and student teachers. On the basis of a recent giftedness definition, we derived a comprehensive set of dimensions of beliefs about the nature of giftedness. The dimensions reflect the content and development of giftedness as discussed in scientific debates. Specifically, we first assessed beliefs about the nature of giftedness with a questionnaire and examined whether the proposed dimensions of beliefs about the nature of giftedness could be found in our sample. In a second step, we analyzed differences in these beliefs between groups of teachers who differed in the amount of time they had spent teaching in general classrooms and in gifted classes. Finally, we were interested in relations between beliefs about the nature of giftedness and beliefs about the malleability of intelligence.

2.1.1 Conceptions of giftedness

Many different conceptions of giftedness (see Sternberg & Davidson, 2005) and approaches for systematizing them (see e.g., Kaufman & Sternberg, 2008; Mönks & Katzko, 2005) have been proposed. In their comprehensive review, Subotnik et al. (2011) categorized the definitions into five perspectives on giftedness: First, giftedness is seen as a high level of general intelligence that enables persons to be successful in many domains (e.g., Hollingworth, 1925; Terman, 1954). Second, giftedness is also seen as high levels of cognitive abilities but is additionally connected with sensitivities that lead to vulnerabilities and emotional fragility (e.g., Mendaglio & Tillier 2006; Webb, 1994). Third, giftedness is understood as an interplay between intelligence and other psychological variables such as creativity and task persistence (e.g., Gagné, 2005; Heller, Perleth, & Lim, 2005; Renzulli, 2005a). Fourth, mostly from a perspective on nonacademic talent development, gifted individuals have to show excellent performance in a specific domain (e.g., Côté, 1999; van Yperen, 2009). Fifth, not ability but deliberate practice in a domain and the number of opportunities to practice are considered crucial to achieve exceptional performance (e.g., Ericsson, 2014).

On the basis of this systematization, which illustrates the many different meanings that have been assigned to the term giftedness, Subotnik et al. (2011, 2012) added a comprehensive conception of giftedness. They integrated the elements of conceptions of giftedness that have reached a relatively high consensus in the field of research. Specifically, Subotnik et al. (2011) defined giftedness as follows:

> Giftedness is the manifestation of performance or production that is clearly at the upper end of the distribution in a talent domain even relative to that of other high-functioning individuals in that domain. Further, giftedness can be viewed as developmental, in that in the beginning stages, potential is the key variable; in later stages, achievement is the measure of giftedness; and in fully developed talents,

> eminence is the basis on which this label is granted. Psychosocial variables play an essential role in the manifestation of giftedness at every developmental stage. Both cognitive and psychosocial variables are malleable and need to be deliberately cultivated. (p. 7)

This conception of giftedness has garnered broad attention (for comments and criticisms, see Plucker & Callahan, 2012), so that we used it as a framework for investigating teachers' beliefs about the nature of giftedness.

2.1.2 Teachers' beliefs about the nature of giftedness

Before discussing teachers' beliefs about the nature of giftedness, we will provide a short introduction about teachers' beliefs in general. In contrast to definitions and scientific conceptions that are explicit theories, which are typically based on or are accessible for empirical examinations (Sternberg, 1985), beliefs as part of individuals' implicit theories do not necessarily have to be in line with empirical results (Baudson & Preckel, 2013a; Sternberg, 1985; Watt & Richardson, 2014). Beliefs can be defined as personal understandings and assumptions about phenomena and objects; such understandings and assumptions are subjectively viewed as true and have both implicit and explicit components (Voss, Kleickmann, Kunter, & Hachfeld, 2013). Fives and Buehl (2012) identified three functions of beliefs: Beliefs filter information and experience, frame situations and problems, and guide intentions and actions. Moreover, (indirect) associations between teachers' beliefs and student outcomes have been found. For example, Watt and Richardson (2014) summarized from the body of research that teachers' beliefs about students' achievement are associated with student variables such as achievement, learning, and perceptions of competence.

In the realm of giftedness, teachers' beliefs seem to play a role in how they view gifted students (e.g., Baudson & Preckel, 2013a; Endepohls-Ulpe & Ruf, 2005), why they identify certain students as gifted (e.g., Siegle, Moore, Mann, & Wilson, 2010; Zhang & Sternberg, 1998), and how they shape the way they teach gifted students (e.g., Eyre et al., 2002; Rambo & McCoach, 2012). Hence, understanding teachers' beliefs is important and can help, for instance, to tailor training programs with respect to identifying and promoting gifted students.

2.1.3. A conception of giftedness as a framework for beliefs about the nature of giftedness

Previous research has accumulated a broad spectrum of insights into teachers' beliefs about the nature of giftedness (e.g., Endepohls-Ulpe & Ruf, 2005; Hany, 1997; Miller, 2009) but still lacks an overall systematization that will enable teachers' beliefs to be integrated into scientific discussions about the meaning of giftedness. To close this gap, a comprehensive conception of giftedness such as the one proposed by Subotnik et al. (2011, 2012) can be used as a framework for extracting crucial facets of the views on giftedness that have divided and unified this research area. Such an approach can be beneficial for obtaining a comprehensive picture of the extent to which teachers' beliefs are aligned with often discussed aspects of scientific conceptions of giftedness.

On the basis of Subotnik et al.'s (2011, 2012) conception, it was possible to derive dimensions reflecting teachers' beliefs about what giftedness is and how it develops. Concerning the question of the meaning of giftedness, four dimensions could be deduced. First, teachers might see giftedness as domain-specific, meaning that gifted students show high ability usually in one or a few domains. Alternatively, teachers might have a holistic view on giftedness that implies that gifted students have to show (the potential for) superior performance in many domains. Second, teachers might also differ in their beliefs about whether giftedness influences students in a similar way so that gifted students are more similar to each other than to other students (i.e., form a homogeneous group) or that gifted students can be very heterogeneous (i.e., form a heterogeneous group). Third, teachers can have different views on whether gifted students have or do not have to show exceptional achievement to be identified as gifted. Fourth, teachers might differ in their views on whether students' intelligence alone would be sufficient to determine giftedness.

Concerning the question of the development of giftedness, an additional four dimensions could be derived from Subotnik et al.'s (2011, 2012) conception of giftedness: Fifth, teachers might believe that a student can develop giftedness or alternatively that giftedness is a trait that a student either has or does not have. Sixth, teachers might agree or disagree with the notion that giftedness is a result of an interplay between many different factors (i.e., biological, pedagogical, psychological, and psychosocial). Seventh, the role of deliberate practice to achieve giftedness can be viewed differently by teachers. Finally, teachers might differ in their beliefs about whether the content of giftedness remains the same across the lifespan or whether the key variables for giftedness are different for students and adults.

2.1.4 Research on teachers' beliefs about the nature of giftedness

When considering the above-mentioned framework, a relatively large body of research is related to teachers' beliefs about the meaning of giftedness. Sternberg and Zhang (1995; Zhang & Sternberg, 1998) illustrated the importance of five criteria when judging whether a student is gifted or not: excellence in one or more domains, rarity of excellent individuals in a corresponding domain or domains, demonstrability of excellence via at least one valid test, productivity shown through a visible outcome, and social valuation of the corresponding domain(s). Endepohls-Ulpe and Ruf (2005) and Hany (1997) showed that German elementary and secondary school teachers believe that the core characteristics of giftedness are attributes related to intelligence and motivation as well as high achievement at school. Similar results were reported by Miller (2009) for American elementary school teachers. Miller concluded that elementary school teachers favor traditional conceptions of giftedness—an inference that was also made by other researchers (e.g., Moon & Brighton, 2008; Speirs Neumeister, Adams, Pierce, Cassady, & Dixon, 2007). Borland (2009) reported that some educators "cling to the giftedness-equals-high-IQ myth" (p. 237), which indicates a one-trait conception of giftedness. However, Schroth and Helfer (2009) showed that the educators in their study (i.e., school teachers, instructors of gifted children, and school administrators) seemed to be confused about what constitutes giftedness when confronted with a list of scientific definitions of giftedness because they accepted virtually all of them.

Lee (1999) found that most teachers had more domain-specific views on giftedness but that some teachers saw giftedness as holistic, meaning that students had to excel in all academic domains to be seen as gifted. Furthermore, the author stated that teachers believe that gifted students' behavior differs from that of their peers. Combined with results from Baudson and Preckel (2013a) who found that teachers believe that gifted students have low social competence and emotional problems, teachers might think that giftedness manifests itself in similar ways across students.

There is little information about teachers' beliefs about whether and how giftedness develops. Reis and Renzulli (2009) and Worrell (2009) reported that giftedness is often seen as a trait that is stable over time. However, the belief that giftedness can be due purely to deliberate practice has also been mentioned (i.e., the 10,000-hr practice rule; e.g., Ericsson, Nandagopal, & Roring, 2005; Hambrick et al., 2014). Furthermore, Sternberg and Zhang (1995) stated that whereas actual achievement and productivity are criteria for labeling adolescents and adults as gifted, children can be considered gifted when they show the potential for high performance.

Taken together, excellence in comparison with peers concerning their cognitive abilities or achievement was consistently seen as component of teachers' beliefs about the nature of giftedness (e.g., Endepohls-Ulpe & Ruf, 2005; Lee, 1999; Miller, 2009; Sternberg & Zhang, 1995). However, the referenced literature has also shown that teachers have diverse beliefs about the nature of giftedness. Although the literature offers valuable insights, it does not provide a comprehensive view on teachers' beliefs about the nature of giftedness that align with often discussed characteristics of scientific conceptions of giftedness.

2.1.5 Differences in teachers' beliefs about the nature of giftedness

Differences in teachers' beliefs about the nature of giftedness might occur on a mean or structural level. Concerning mean-level differences, teachers' decisions about which students they view as gifted have been found to be influenced by student characteristics: Girls have been found to have lower chances of being identified as gifted than boys (e.g., Bianco, Harris, Garrison-Wade, & Leech, 2011; Endepohls-Ulpe & Ruf, 2005; Hernández-Torrano, Prieto, Ferrándiz, Bermejo, & Sáinz, 2013) as well as students from families with lower socioeconomic status or students with certain ethnic backgrounds (e.g., McBee, 2006; Speirs Neumeister et al., 2007). With regard to teacher variables, previous research has shown that differences in teachers' beliefs about the nature of giftedness are related to teachers' level of experience with giftedness, for instance, through training in the area of giftedness (Copenhaver & Intyre, 1992; Siegle & Powell, 2004) or teaching experience with gifted students (Endepohls-Ulpe & Ruf, 2005; Schack & Starko, 1990). Concerning teaching experience in general classrooms, Guskin, Peng, and Simon (1992) found that when teachers were asked to rate students from simulated profiles with embedded characteristics reflecting giftedness, school teachers rated such students as more able than student teachers did. However, concerning beliefs and attitudes about gifted education (Bégin & Gagné, 1994; Chessman, 2010; Cramond & Martin, 1987) and the personality of giftedness (Baudson & Preckel, 2013a), teaching experience in general classrooms did not have an impact.

Differences in beliefs about the nature of giftedness can also occur on a structural level, but relevant studies are still rather rare. Indicators of different belief structures in the cross-cultural context have been found: for example, only three of the five factors of the Pentagonal Implicit Theory of Giftedness were found for student teachers in China (Zhang & Hui, 2002). Also, when American and German teachers were presented with a list of possible characteristics of gifted students, different factor structures were found (Busse, Dahme, Wagner, & Wieczerkowski, 1986). The factor structure for American teachers consisted of five factors, whereas the structure for German teachers consisted of seven.

Studies in one society have shown more mixed results: García-Cepero and McCoach (2009) found only partial measurement invariance in beliefs about intelligence and beliefs about the identification of gifted students between American school teachers and professors from schools of education. Also, in a qualitative study, Goodnough (2000) found that students who took a course on giftedness at a university developed broad multidimensional beliefs about the meaning of giftedness. However, Miller's (2009) school teachers who had different levels of postbaccalaureate training in the area of gifted education did not differ, for example, in the complexity of their beliefs about the nature of giftedness.

2.1.6 Associations between beliefs about giftedness and beliefs about intelligence

Intelligence is central to most definition of giftedness (Subotnik et al., 2011, 2012). However, little is known about the question of whether (or how) beliefs about the malleability of intelligence are associated with beliefs about the nature of giftedness. According to Dweck and colleagues (e.g., Dweck, Chiu, & Hong, 1995), people tend to embrace either an entity theory of intelligence (i.e., intelligence is a stable trait that cannot be changed) or an incremental theory of intelligence (i.e., intelligence is malleable and changeable). Jones, Bryant, Snyder, and Malone (2012) found that about three quarters of student teachers and school teachers viewed intelligence as rather incremental. Jonsson, Beach, Korp, and Erlandson (2012) differentiated between teachers' subjects and found that teachers of languages, social sciences, or practical subjects had a stronger preference for an incremental theory than math teachers did. In Rattan, Good, and Dweck's (2012) study, teachers embracing an entity theory of intelligence had more of a tendency to infer low ability to low-achieving students than teachers who embraced an incremental theory did. These low-achieving students perceived teachers' intelligence-related beliefs about them and reacted with lower motivation.

To the best of our knowledge, beliefs about the malleability of intelligence have not yet been explicitly related to beliefs about the nature of giftedness. García-Cepero and

McCoach (2009) found no link between teachers' beliefs about the malleability of intelligence and their preferences for either a method based solely on intelligence tests versus a method that used multiple criteria to identify gifted students. However, other direct associations (e.g., with the belief that giftedness is malleable) are possible.

2.2 Research Questions

Teachers' beliefs can be important with respect to the identification (e.g., Endepohls-Ulpe & Ruf, 2005; Siegle et al., 2010; Sternberg & Zhang, 1995) and promotion of gifted students (e.g., Eyre et al., 2002; Rambo & McCoach, 2012). Research has shown that beliefs about the nature of giftedness can differ between groups of teachers—on a mean level (e.g., Endepohls-Ulpe & Ruf, 2005; Guskin et al., 1992) or on a structural level (e.g., Busse et al., 1986; García-Cepero & McCoach, 2009). The present study aimed to add to the body of research by presenting a systematically derived set of dimensions of beliefs about the content and development of giftedness that was based on a scientific conception of giftedness (i.e., the one proposed by Subotnik et al., 2011, 2012). These dimension were chosen to capture different teachers' beliefs that are in alignment with often discussed aspects of the meaning of giftedness.

The study was conducted in the context of an enrichment program called the Hector Children's Academy Program (HCAP). The HCAP has the goal of providing classes for the 10% most gifted students in Baden-Württemberg, Germany. The classes are voluntary and are held outside of the regular school curriculum. The HCAP has 60 academies that are typically located at one or more elementary schools (for further information about the HCAP, please see Rothenbusch et al., 2016). In the present study, we drew on data from four different groups of teachers (i.e., elementary school teachers, two groups of teachers from the HCAP, and student teachers) and investigated the following research questions:

First, how are beliefs about the nature of giftedness structured? We investigated the structure of these beliefs in the total sample. Because we proposed that beliefs about the nature of giftedness would be represented by theoretically distinct dimensions, we expected low correlations between the different dimensions.

Second, will the (a) structure and (b) mean level of the beliefs about the nature of giftedness differ between the different groups of teachers? To address the structure, we investigated the extent to which the structure reflected invariance across the different

Table 2.1

Groups With and Without Teaching Experience in Gifted Classes and General Classrooms

	EST	HST	HI	ST
Teaching experience in gifted classes	No	Yes	Yes	No
Teaching experience in general classrooms	Yes	Yes	No	No

Note. EST = elementary school teachers; HST = HCAP school teachers, HI = HCAP instructors; ST = student teachers.

teacher samples. Due to the rather inconclusive state of research on measurement invariance in teachers' beliefs about the nature of giftedness in one society (see García-Cepero & McCoach, 2009; Goodnough, 2000; Miller, 2009), we did not know whether we should expect measurement invariance. To address the mean level, we analyzed differences between the groups in their latent mean levels of beliefs about the nature of giftedness. The four groups of teachers that participated in our study differed in whether they had or did not have teaching experience in gifted classes and in general classrooms (see Table 2.1). Previous research has found differences in beliefs about the nature of giftedness with respect to both kinds of teaching experience (Endepohls-Ulpe & Ruf, 2005; Guskin et al., 1992; Schack & Starko, 1990). Therefore, we expected to find differences between the groups in their beliefs about the nature of giftedness. As student teachers lack both kinds of teaching experience, we hypothesized that their beliefs about the nature of giftedness would show the largest differences from the other groups. Furthermore, in the groups with teaching experience (i.e., all except student teachers), we explored whether the number of years spent teaching in gifted classes or general classrooms was related to these teachers' beliefs about the nature of giftedness.

Third, how are beliefs about the nature of giftedness related to beliefs about the malleability of intelligence? Because intelligence is a main component of most giftedness definitions (see Sternberg & Davidson, 2005), we explored the relation between beliefs about the nature of giftedness and beliefs about the malleability of intelligence in two groups of teachers (i.e., elementary school teachers and student teachers). More precisely, we expected that when intelligence was seen as malleable, giftedness would also be perceived as malleable.

2.3 Method

2.3.1 Participants and procedure

A total of 1,235 teachers and instructors from Baden-Württemberg, Germany, participated in our study. They belonged to four groups: (a) 131 elementary school teachers (*M* = 46.42 years, *SD* = 11.53; 85.50% female), (b) 212 school teachers who also taught in the HCAP (*M* = 48.79 years, *SD* = 12.92; 60.66% female), (c) 363 HCAP instructors who were not school teachers but experts in the area they taught in the program (*M* = 48.54 years, *SD* = 12.26; 64.67% female), and (d) 529 student teachers (*M* = 21.68 years, *SD* = 2.26; 64.45% female).

For the present investigation, we combined data from three points of data collection in 2012 and 2013. First, elementary school teachers were asked to answer a questionnaire in the context of a large HCAP effectiveness study. For the effectiveness study, schools were randomly selected out of a pool of schools that had sent students to the HCAP in previous years. Second, all HCAP school teachers and instructors were invited to answer a questionnaire; 80.36% of the HCAP school teachers and 30.44% of the HCAP instructors responded. Data were gathered from HCAP school teachers and instructors from 55 of the 60 participating academies. For elementary school teachers and HCAP school teachers and instructors, participation was voluntary. Third, student teachers attending a university course on pedagogical psychology worked on an online assessment (including our questionnaire) as part of their course. It is important to note that the online assessment took place before the student teachers had taken classes on intelligence or giftedness.

2.3.2 Measures

2.3.2.1 Beliefs about the nature of giftedness

We measured beliefs about the nature of giftedness with a newly developed questionnaire. Using the review on conceptions of giftedness conducted by Subotnik et al. (2011, 2012) and their proposed comprehensive conception of giftedness to guide us, we extracted eight dimensions reflecting the content and development of giftedness as discussed in the scientific field. We generated items covering each of the eight extracted dimensions (see Table 2.2): *Domain-Specific vs. Holistic Giftedness (DHG), Heterogeneity vs.Homogeneity (HH)*, *Importance of Achievement (IA), Importance*

Table 2.2

Dimensions of Beliefs About the Nature of Giftedness

Key characteristics of Subotnik et al.'s (2011, 2012) giftedness definition	Dimensions of beliefs about the nature of giftedness	Scale values
Content of giftedness		
Giftedness is domain-related.	Domain-Specific vs. Holistic Giftedness	Low: Giftedness is domain-specific. High: Giftedness is holistic.
Gifted persons are rather heterogeneous but have intensity, persistence, and ability in common.	Heterogeneity vs. Homogeneity	Low: Gifted students do not form a homogeneous group. High: Gifted students form a homogeneous group.
Giftedness manifests itself during the early stages in the potential to achieve but typically in performance later on.	Importance of Achievement	Low: Giftedness does not need to be shown by superior achievement. High: Giftedness has to be shown by superior achievement.
High intelligence is a necessary but not sufficient condition for giftedness.	Importance of Intelligence	Low: Intelligence is not the most important characteristic of giftedness. High: Intelligence is the most important characteristic of giftedness.
Development of giftedness		
Giftedness can be seen as a developmental process.	Mutable vs. Fixed Giftedness	Low: Giftedness is mutable. High: Giftedness is fixed.
The development of giftedness depends on a conglomerate of biological, pedagogical, psychological, and psychosocial factors.	Interplay of Personal and Environmental Factors	Low: Giftedness does not develop through an interplay of personal and environmental factors. High: Giftedness develops through an interplay of personal and environmental factors.
Giftedness is multifactorial and, therefore, more than the result of deliberate practice.	Influence of Deliberate Practice	Low: Giftedness cannot be the result of deliberate practice. High: Giftedness can be the result of deliberate practice.
The key variables differ in accordance with the developmental stage the individual is in.	Different Key Variables for Children and Adults	Low: The key variables for gifted children and adults are the same. High: The key variables for gifted children and adults are different.

of Intelligence (II), *Mutable vs. Fixed Giftedness (MFG)*, *Interplay of Personal and Environmental Factors (PEF)*, *Influence of Deliberate Practice (IDP)*, and *Different Key Variables for Children and Adults (DKV)*.

The items were revised several times. After creating an item pool comprised of 72 items, our team inspected and piloted it in an online survey administered to 226 student teachers who were not part of the main study. On the basis of the results, we shortened the questionnaire by deleting 32 items, rephrased 11 items, and replaced 1 item. After another inspection of the items by our team and two teachers, only a few changes in wording were made. The resulting 40-item questionnaire was administered to our sample. Participants indicated their agreement with the 40 items on a 4-point Likert-type scale ranging from 1 (*disagree*) to 4 (*agree*). Results of the inspection of the factor structure will be presented in the Result section. Factor determinacies that were based on the entire sample and that are indices for the reliabilities of the dimensions can be found in Table 2.3.

2.3.2.2 Beliefs about the malleability of intelligence

To measure teachers' beliefs about the malleability of intelligence, we used a translated version of the "Theories of Intelligence Scale – Self Form for Adults" (Dweck, 1999). This scale consists of eight items that have to be answered on a 6-point Likert-type scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The scale captures two opposite strands of intelligence beliefs: The low end represents an entity theory of intelligence, whereas the high end represents an incremental theory of intelligence. The four items that measured agreement with an entity theory had to be recoded. An example item is "You can always greatly change how intelligent you are." Due to practical issues with the HCAP, we could only administer this scale to the student teachers (M = 3.81, SD = 0.97) and elementary school teachers (M = 3.61, SD = 0.99) but not to HCAP school teachers and instructors. The internal consistency estimates for the group of elementary school teachers and the group of student teachers were α = .92 and α = .93, respectively.

Furthermore, with one item each, we assessed the number of years spent teaching in gifted classes (HCAP school teachers: M = 2.74 years, SD = 1.70; HCAP instructors: M = 2.45 years, SD = 2.26) and the number of years spent teaching in general classrooms (elementary school teachers: M = 18.51 years, SD = 11.89; HCAP school teachers: M = 17.87 years, SD = 12.21).

2.3.3 Analyses

We ran the analyses in the Mplus 7.11 software package (Muthén & Muthén, 1998-2015). We used z-standardized data but retained the original metric of the dichotomous variables. Missing data ranged from 0.65% for teachers' responses to item IDP1 to 27.41% for the number of years spent teaching in general classrooms. We applied the full information maximum likelihood algorithm, which estimates missing values on the basis of the full information from the covariance matrices (Enders, 2010). The use of multiple significance tests inflates the Type I error rate, meaning that it increases the chance of falsely rejecting the null hypothesis. To adjust for the false discovery rate, we used the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995).

2.3.3.1 Dimensionality of beliefs about the nature of giftedness

In order to inspect the factor structure and psychometric quality of the newly developed questionnaire for measuring teachers' beliefs about the nature of giftedness, we inspected the items' skewness and kurtosis. According to George and Mallery (2012), values between +/-2.00 for both indices are acceptable in most cases. After the item inspection, we applied exploratory structural equation modeling (ESEM). ESEM integrates EFA, confirmatory factor analysis (CFA), and structural equation modeling (SEM) without the restrictive assumption required in CFA that there can be no cross-loadings (Marsh, Morin, Parker, & Kaur, 2014). ESEM overcomes some additional limitations of EFAs (Asparouhov & Muthén, 2009; Marsh, Nagengast, & Morin, 2012). For example, ESEM allows measurement invariance testing. We used an oblique geomin rotation with an epsilon value of .5 as recommended by Marsh et al. (2010) and Marsh et al. (2012). We assessed the model fit according to the fit indices computed for these analyses (see the passage about "goodness of fit" at the end of the Analyses section).

Cudeck (2000) and Tinsley and Tinsley (1987) stated that a common rule of thumb for selecting items is to choose items with factor loadings of .30 or higher. Thus, we decided to select items with factor loadings equal to or larger than .30 on one factor and with no cross-loadings equal to or larger than .30 on other factors. Associations between the dimensions of beliefs about the nature of giftedness were examined by correlating each dimension's latent factor with each other.

Table 2.3

Factor Loadings From the ESEM Solution Based on Responses to 30 Items

	DHG	HH	IA	II	IMG	PEF	IDP	DKV	R^2	Items
F1 (DHG)										*Domain-Specific vs. Holistic Giftedness (DHG)*
DHG1R	**.35**	.08	.01	-.12	-.01	-.20	.15	-. 06	.23	Gifted children can be average or below average in some areas in comparison with other children of the same age.
DHG2	**.54**	.13	-.03	.15	.11	.01	.10	. 00	.43	Theoretically, a gifted child has a great deal of potential to achieve in nearly all academic areas.
DHG3	**.57**	.03	.26	.08	.05	.07	-.03	. 03	.50	Gifted children are good at most school subjects.
DHG4	**.80**	.03	.04	.07	.00	.04	-.03	-. 01	.69	Gifted children are at the top of the class in almost all ability areas.
F2 (HH)										*Heterogeneity vs. Homogeneity of Gifted Children (HH)*
HH1	-.11	**.31**	.06	.01	.19	.06	-.04	. 13	.15	Gifted children have different social and emotional needs from other students.
HH2	.08	**.62**	.09	-.06	-.03	-.02	.03	-. 05	.43	Gifted children often have similar beliefs and attitudes.
HH3	.03	**.63**	.08	.09	-.09	.01	-.01	-. 04	.44	Gifted children are more similar to each other in comparison with other children.
HH4	.04	**.70**	-.07	.07	.05	.01	.01	. 01	.53	Gifted children often have a similar view on the world.
F3 (IA)										*Importance of Achievement (IA)*
IA1	.04	-.03	**.35**	.13	-.01	.07	.02	-. 02	.16	Giftedness involves superior achievement in comparison with same-age children.
IA2	-.02	.02	**.58**	.08	.04	-.01	.01	. 01	.34	Children's giftedness must show itself in exceptionally high performance in certain domains.
IA3	.09	.02	**.46**	-.19	-.04	-.05	.12	. 00	.32	A child's intelligence score might be very high; however, if the child does not show better achievement then same-age children, she/he is not gifted.
IA4R	.18	.09	**.32**	-.28	.03	-.07	.02	-. 08	.26	Even though children might not show exceptional performance, they might be gifted.
F4 (II)										*Importance of Intelligence (II)*
II1	.09	.08	.03	**.62**	.11	-.04	.01	-. 05	.49	Ultimately, the IQ score separates gifted from not gifted students.
II2	.12	.02	-.01	**.64**	.02	-.05	-.06	-. 02	.48	The most important characteristic of gifted children is their high intelligence.
F5 (MFG)										*Mutable vs. Fixed Giftedness (MFG)*
MFG1	-.01	.04	.14	.09	**.48**	-.11	.02	-. 06	.34	Giftedness is stable across time and independent of experience.
MFG2	.04	-.02	.02	.12	**.76**	.00	-.05	. 01	.67	Gifted once means gifted always.
MFG3R	.06	-.04	-.08	-.01	**.57**	-.11	-.05	-. 19	.49	Individuals who are identified as gifted as children might not necessarily be gifted as adults.

(continued)

Table 2.3 (continued)

Factor Loadings From the ESEM Solution Based on Responses to 30 Items

	DHG	HH	IA	II	IMG	PEF	IDP	DKV	R^2	Items
F6 (PEF)										*Interplay of Personal and Environmental Factors (PEF)*
PEF1	.08	.07	-.06	.06	-.02	**.54**	.07	-.02	.32	Sociocultural conditioning is of crucial importance to any kind of development in a child and therefore also for the development of giftedness.
PEF2	.06	-.03	.07	-.05	.00	**.56**	-.07	.02	.32	Giftedness develops through an interplay of biological, pedagogical, psychological, and psychosocial factors.
PEF3	.01	.02	-.06	-.08	-.11	**.63**	.09	.01	.48	Giftedness develops not only through personal but also through environmental factors.
PEF4	-.02	-.06	.15	.04	-.24	**.43**	.05	.04	.35	Two children can be genetically very similar but do not both have to be gifted due to environmental influences.
F7 (IDP)										*Influence of Deliberate Practice (IDP)*
IDP1	-.04	-.04	.10	.01	-.09	.12	**.54**	-.01	.37	Almost every academic ability can be trained, therefore, giftedness too.
IDP2	.03	.00	.01	-.02	-.05	.03	**.75**	-.03	.59	It almost does not matter which preconditions a child has. With intensive promotion, every child can be gifted.
IDP3	-.06	.06	.06	-.07	-.04	.05	**.60**	.02	.42	Intensive training produces giftedness, not the intelligence of a child.
IDP4	.08	.02	.08	-.10	-.02	-.03	**.66**	.06	.52	Independent of genetic dispositions, a child can be gifted with sufficient practice.
F8 (DKV)										*Different Key Variables for Children and Adults (DKV)*
DKV1	-.08	.10	.01	.03	.07	.12	.04	**.50**	.29	Giftedness manifests itself differently in children than in adults.
DKV2R	.00	-.05	-.22	-.14	-.03	.00	.04	**.63**	.52	A label of gifted means the same thing for children and adults.
DKV3	-.09	.06	.07	.14	.06	.20	-.08	**.37**	.23	Expectations of gifted children differ from those of gifted adults.
DKV4R	.01	-.06	-.02	-.23	-.20	-.04	.00	**.58**	.54	The same criteria apply to both children and adults in identifying them as gifted.
DKV5	.01	-.03	.10	.06	-.17	.04	-.01	**.70**	.58	When discussing giftedness, you have to differentiate between children and adults.
Factor determinacy	.89	.85	.79	.83	.88	.83	.88	.88		

Note. Factor loadings ≥ .30 are in bold. N = 1,235. ESEM = exploratory structural equation modeling; R (as in DHG1R) = reverse-scored.

2.3.3.2 Group comparisons

The first step for group comparisons was to test for measurement invariance across groups. As suggested by Marsh et al. (2012), we tested different multiple-group models by applying the ESEM approach. First, we estimated a four-group ESEM with all factor loadings and intercepts varying freely in all groups to test for configural measurement invariance. Finding no configural measurement invariance would indicate that there was not a similar factor pattern across the groups. Second, we restricted the factor loadings to be invariant across the groups to test for weak measurement invariance. A lack of invariance in the factor loadings might indicate that the constructs were interpreted differently by the different groups. When weak invariance fails, partial weak measurement invariance can be established. Then ESEM-within-CFA (EWC) analyses has to be conducted because the partial releasing of factor loadings is not allowed in the ESEM approach (Marsh et al., 2014). Third, we tested for strong measurement invariance by additionally restricting the item intercepts to be invariant across the groups. Strong measurement invariance is needed to compare the latent means between groups. If strong measurement invariance cannot be achieved, testing for partial measurement invariance can be conducted by freeing individual item intercepts within the ESEM or the EWC approach. Following Chen's (2008) recommendation, results from models with partial strong measurement invariance should be compared with the results from models with full strong measurement invariance to inspect the impact of noninvariance on the results. Small differences might (but do not have to) indicate that group comparisons are justifiable. Therefore, we tested the parameter estimates for the partial and full strong measurement invariance models for significant differences.

In a second step, if—at least partial—strong measurement invariance can be achieved, latent means can be compared between groups. We estimated the latent mean differences and tested whether they were significantly different from 0. For each dimension, we tested for differences between the four groups of teachers.

2.3.3.3 Associations with beliefs about the malleability of intelligence and years spent teaching in general and gifted classes

To examine how beliefs about the nature of giftedness are associated with beliefs about the malleability of intelligence, years spent teaching in general classrooms, and years spent teaching in the HCAP, we calculated regressions. To represent beliefs about the nature of giftedness, we used the factor scores from the model of partial strong measurement invariance as the dependent variables. The manifest variables *beliefs about* *the malleability of intelligence*, *years spent teaching in general classrooms*, and *years spent teaching in gifted classes* were the independent variables.

2.3.3.4 Goodness of fit

Model fit was assessed with the χ^2 goodness-of-fit statistic and the following descriptive indices (Hu & Bentler, 1999). We computed the Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI). For both the TLI and CFI, values above .90 indicate an acceptable fit and values above .95 an excellent fit to the data. Furthermore, we used the Root Mean Square Error of Approximation (RMSEA; values below .05 are considered to indicate a close fit to the data) and the Standardized Root Mean Square Residual (SRMR; values below .08 are considered to indicate a good model fit).

To compare the relative fit between nested models (i.e., a model with the same parameters as another but with some additional restrictions) in the analysis of measurement invariance between groups, we used the guidelines proposed by Chen (2007): The CFI, RMSEA, and SRMR values should not change more than -.010, .015, and .030, respectively, when measuring invariance in the loadings (i.e., weak measurement invariance) and not more than -.010, .015, and .010, respectively, when measuring invariance in the intercepts (i.e., strong measurement invariance).

2.4 Results

2.4.1 How are beliefs about the nature of giftedness structured?

The first aim of our investigation was to analyze whether the dimensions that were proposed for representing beliefs about the nature of giftedness could be found in a large, diverse sample of teachers (i.e., elementary school teachers, HCAP school teachers, HCAP instructors, and student teachers). On a descriptive level, the skewness and kurtosis values ranged between +/-2.00. In order to investigate the structure of beliefs about the nature of giftedness in our sample, we computed an ESEM analysis. The model results showed acceptable values, $\chi^2(488) = 1148.189$, RMSEA = .033, CFI = .939, TLI = .903, SRMR = .022. However, a closer inspection showed that some items had low factor loadings on the target factors or high cross-loadings. Several additional ESEM analyses with successively fewer items were computed until we arrived at a satisfactory factor solution that represented the intended dimensions of beliefs about the nature of giftedness. In this process, we had to exclude 10 of the 40 items (5 items had too low factor loadings,

Table 2.4

Latent Factor Correlations Between the Dimensions of Beliefs About the Nature of Giftedness

	DHG	HH	IA	II	MFG	PEF	IDP
Domain-Specific vs. Holistic Giftedness (DHG)	-						
Heterogeneity vs. Homogeneity of Gifted Children (HH)	.19***	-					
Importance of Achievement (IA)	.25***	.13***	-				
Importance of Intelligence (II)	.16***	.15***	.01	-			
Mutable vs. Fixed Giftedness (MFG)	.11***	.08***	.00	.29***	-		
Interplay of Personal and Environmental Factors (PEF)	-.02	.02	.06	.05	-.25***	-	
Influence of Deliberate Practice (IDP)	.09***	.06**	.18***	-.19***	-.18***	.10***	-
Different Key Variables for Children and Adults	-.11***	-.02	-.04	-.10***	-.25***	.19***	.02

Note. $N = 1{,}235$. Significance tests were set to an overall level of $\alpha = .05$ and adjusted with the Benjamini-Hochberg procedure (1995).
** $p < .01$. *** $p < .001$.

1 item had a too high cross-loading, and 4 items had target-factor loadings that were too low and cross-loadings that were too high). Further information about the excluded items can be found in the Appendix.

The resulting 30-item solution showed good fit indices, $\chi^2(223) = 446.906$, RMSEA = .029, CFI = .973, TLI = .947, SRMR = .017, with substantial factor loadings (see Table 2.3). As can be seen in Table 2.3, the third factor *Importance of Achievement* had rather low factor loadings overall, and the fourth factor *Importance of Intelligence* consisted of only two items with high factor loadings. We checked whether the removal of these two factors would lead to changes in the factor structure. A six-factor solution showed that the remaining items still had high loadings on their target factors and low loadings on the remaining factors, and the model fit was descriptively slightly worse, $\chi^2(147) = 333.998$, RMSEA = .032, CFI = .972, TLI = .948, SRMR = .018. Because these factors are theoretically important, we decided to keep them. However, results concerning these factors have to be viewed with some caution.

The correlations between the dimensions were relatively low (see Table 2.4), thereby indicating that the dimensions were relatively independent of each other ($.00 \leq r \leq |.29|$). *Domain-Specific vs. Holistic Giftedness*, *Mutable vs. Fixed Giftedness*, and *Influence of Deliberate Practice* had the largest number of six significant correlations and *Importance of Achievement* had the lowest number of three significant associations with other dimensions of beliefs about giftedness.

2.4.2 Do teachers differ in their beliefs about the nature of giftedness?

2.4.2.1 Differences in the structure of beliefs about the nature of giftedness

We computed measurement invariance tests to determine whether different groups had similar interpretations of the items and factors. Only if measurement invariance—or at least partial measurement invariance—could be established would comparisons between groups have the potential to be fair. As shown in Table 2.5, the four-group ESEM in which all factor loadings and intercepts varied freely showed a good fit to the data. Next, we fixed the factor loadings to invariance across the four groups. The CFI value was too low, so we could not establish weak measurement invariance. However, by freeing some of the factor loadings, partial weak invariance could be established. To be able to free the factor loadings, we had to change from the ESEM model to the ESEM-within-CFA framework as recommended by Marsh et al. (2014). We freed the factor loadings with the highest deviations according to the modification indices until the model fit indices were within the recommended acceptable range of change in model fit. Although the factor loadings of 19 items had to be freed in one or more groups of teachers, the mapping of items onto their target factors (factor loadings $\geq .30$) remained the same in all groups. The HCAP instructors had the smallest number and the student teachers had the largest number of noninvariant factor loadings (i.e., 8 and 15, respectively). To test for partial strong measurement invariance, we set the intercepts to invariance across the groups. The change in fit indices was within the acceptable range. Therefore, partial strong measurement invariance could be established. As recommended by Chen (2007), we compared the results of the full and partial strong measurement models for significant differences. Two correlations differed significantly between the two four-group models.

In the group of HCAP instructors, the correlation between *Importance of Achievement* and *Importance of Intelligence* was more negative in the partial strong measurement model (full strong measurement invariance: $r = -.16$, $p > .05$; partial strong measurement invariance: $r = -.40$, $p < .05$). In the group of student teachers, the correlation between *Importance of Achievement* and *Domain-Specific vs. Holistic Giftedness* were nonsignificant in both models but differed significantly from each other (full strong measurement invariance: $r = .12$, $p > .05$; partial strong measurement

Table 2.5

Series of Models Used to Test for Measurement Invariance Between the Four Groups of Teachers: Elementary School Teachers, HCAP School Teachers, HCAP Instructors, and Student Teachers

Model	MI level	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Δ CFI	Δ RMSEA	Δ SRMR
ESEM	Configural	1236.909	892	.959	.920	.035	.027			
ESEM	Weak	2070.052	1420	.923	.905	.039	.049	-.036	.004	.022
EWC	Partial weak	1788.227	1374	.951	.938	.031	.042	-.008	-.004	.015
EWC	Partial strong	1920.507	1441	.943	.931	.033	.044	-.008	.002	.002
EWC	Partial strong + invariant factor means	2174.658	1465	.916	.900	.040	.058	-.027	.007	.014

Note. ESEM = exploratory structural equation modeling; EWC = ESEM within CFA; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation; Δ = Difference.

Table 2.6

Manifest Means of the Dimensions of the Beliefs About the Nature of Giftedness

	All teachers		Elementary school teachers		HCAP school teachers		HCAP instructors		Student teachers	
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
Domain-Specific vs. Holistic Giftedness	2.00	(0.56)	1.99	(0.57)	1.98	(0.59)	2.02	(0.51)	1.99	(0.57)
Heterogeneity vs. Homogeneity	2.19	(0.56)	2.22	(0.52)	2.27	(0.53)	2.26	(0.56)	2.12	(0.57)
Importance of Achievement	2.24	(0.54)	2.22	(0.53)	2.14	(0.53)	2.21	(0.52)	2.30	(0.55)
Importance of Intelligence	2.72	(0.72)	2.89	(0.69)	2.85	(0.72)	2.80	(0.73)	2.57	(0.71)
Mutable vs. Fixed Giftedness	2.37	(0.69)	2.59	(0.65)	2.45	(0.69)	2.33	(0.74)	2.31	(0.65)
Interplay of Personal and Environmental Factors	2.97	(0.56)	3.03	(0.53)	2.93	(0.54)	2.98	(0.62)	2.96	(0.53)
Influence of Deliberate Practice	1.64	(0.55)	1.39	(0.39)	1.45	(0.44)	1.57	(0.51)	1.83	(0.58)
Different Key Variables for Children and Adults	2.91	(0.56)	2.84	(0.53)	2.80	(0.59)	2.85	(0.57)	3.00	(0.53)

Note. *Min* = 1, *Max* = 4. For each factor, the unweighted manifest mean of the items that loaded ≥ .30 on the factor is depicted.

invariance: $r = -.03$, $p > .05$). With partial measurement invariance, we inspected the latent mean differences between the groups of teachers. However, although the latent mean values did not differ significantly between the partial and full strong measurement invariance models, interpretations had to be made cautiously because structural differences between the groups might still be (at least partly) responsible for the mean-level difference.

2.4.2.2 Differences in mean levels of beliefs about the nature of giftedness

On a descriptive level, Table 2.6 shows the manifest means of the factor scores (calculated by computing the mean of each participant's unweighted items that loaded higher than .30 on a factor) for the total sample and the four groups of teachers. The lowest scale value occurred on *Influence of Deliberate Practice* ($M = 1.64$, $SD = 0.55$), showing that teachers did not tend to believe that giftedness is a product of deliberate practice. *Interplay of Personal and Environmental Factors* ($M = 2.97$, $SD = 0.56$) had the highest value, indicating that teachers tended to believe that giftedness develops through an interplay of personal and environmental factors.

To test for differences in latent means between the groups, we additionally restricted the latent scale means to be invariant in the model with partial strong measurement invariance. The change of model fit was outside the acceptable range (see Table 2.5). Therefore, we could conclude that there were mean differences between the groups. The six comparisons that could be made on a latent mean level for each of the eight dimensions (see Table 2.7) showed that student teachers differed the most from the other groups of teachers (i.e., on four dimensions), especially from the group of elementary school teachers (i.e., on seven dimensions). In contrast to the other three groups of teachers, student teachers saw gifted children as more heterogeneous, did not believe as much as the other groups that intelligence is the most important factor for giftedness, did not disagree as much as the other groups that giftedness can be achieved only through deliberate practice, and were more likely to agree that the key variables for giftedness differ between children and adults. Furthermore, student teachers saw giftedness as more strongly connected to superior achievement than elementary school teachers and HCAP school teachers did and were more likely than elementary school teachers to agree that giftedness is malleable and develops through an interplay of different factors.

Table 2.7

Differences in Latent Means Between Teacher Groups With Different Levels of Teaching Experience With Gifted Classes and General Classrooms

			DHG		HH		IA		II		MFG		PEF		IDP		DKV	
			Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)	Δ *M*	(*SE*)
HST	vs.	EST	0.03	(0.13)	-0.01	(0.13)	-0.09	(0.14)	0.14	(0.12)	-0.28	(0.16)	-0.24	(0.12)	0.14	(0.14)	-0.22	(0.13)
HI	vs.	EST	0.12	(0.11)	-0.03	(0.11)	0.09	(0.13)	0.07	(0.12)	-0.51**	(0.16)	-0.28	(0.13)	0.54***	(0.15)	-0.09	(0.11)
ST	vs.	EST	0.03	(0.11)	-0.34**	(0.11)	0.34*	(0.13)	-0.31*	(0.12)	-0.49**	(0.15)	-0.37**	(0.12)	1.36***	(0.18)	0.32*	(0.13)
ST	vs.	HST	0.00	(0.10)	-0.33**	(0.12)	0.43**	(0.14)	-0.45***	(0.12)	-0.20	(0.13)	-0.13	(0.12)	1.22***	(0.17)	0.54***	(0.12)
ST	vs.	HI	-0.09	(0.08)	-0.31**	(0.09)	0.25	(0.11)	-0.38***	(0.11)	0.02	(0.11)	-0.09	(0.11)	0.81***	(0.14)	0.40***	(0.10)
HI	vs.	HST	0.09	(0.10)	-0.02	(0.11)	0.18	(0.12)	-0.07	(0.11)	-0.22	(0.14)	-0.04	(0.12)	0.40**	(0.13)	0.14	(0.11)

Note. EST = elementary school teachers; HST = HCAP school teachers, HI = HCAP instructors; ST = Student teachers; Δ = Difference; DHG = Domain-Specific vs. Holistic Giftedness; HH = Heterogeneity vs. Homogeneity of Gifted Children; IA = Importance of Achievement; II = Importance of Intelligence; MFG = Mutable vs. Fixed Giftedness; PEF = Interplay of Personal and Environmental Factors; IDP = Influence of Deliberate Practice; DKV = Different Key Variables for Children and Adults. A positive mean difference indicates a higher level of the mean score of the first group and a negative mean difference shows that the mean value of the second group is higher. Significance tests were set to an overall level of $\alpha = .05$ and adjusted with the Benjamini-Hochberg procedure (1995).

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

The other three groups of teachers' beliefs about the nature of giftedness tended to be similar to each other. Only three out of 24 comparison showed differences in beliefs: HCAP school teachers' beliefs did not differ significantly from the beliefs of elementary school teachers. HCAP instructors were more likely than elementary school teachers to agree that giftedness is malleable and were less likely than elementary school teachers and HCAP school teachers to disagree that giftedness can be achieved through deliberate practice.

We furthermore analyzed whether the time spent teaching gifted classes or general classes was associated with beliefs about the nature of giftedness (see Table 2.8). Therefore, we computed regression analyses with the factor scores of the dimensions of beliefs about the nature of giftedness as the dependent variables and the two duration variables as independent variables. Concerning the two groups of HCAP teachers, the time spent teaching gifted classes was associated with two dimensions of beliefs about the nature of giftedness in the group of HCAP instructors but with none of the dimensions in the group of HCAP school teachers. The more time HCAP instructors spent teaching gifted classes, the more they believed that giftedness develops through an interplay of diverse factors and the more they believed that students can develop giftedness through deliberate practice. Only the latter association was significantly different between the two groups ($b_{\text{HCAP school teachers}} - b_{\text{HCAP instructors}} = -0.28$, $p < .05$).

For the two groups of school teachers (see Table 2.8), we did not find significant associations between the time spent teaching in general classrooms and beliefs about the nature of giftedness in the group of elementary school teachers or in the group of HCAP school teachers. Furthermore, the associations were not significantly different between the groups (all ps > .05).

2.4.3 How are beliefs about the nature of giftedness related to beliefs about the malleability of intelligence?

Beliefs about the malleability of intelligence were associated with several dimensions of beliefs about the nature of giftedness (see Table 2.8). The more elementary school teachers and student teachers believed in an incremental theory of intelligence, the less they believed that intelligence is important for giftedness or that giftedness is immutable. The more both groups saw intelligence as incremental, the more they believed

Table 2.8

Associations Between Beliefs About the Nature of Giftedness and Years Spent Teaching in Gifted Classes, Years Spent Teaching in General Classrooms, and Beliefs About the Malleability of Intelligence

		DHG		HH		IA		II		MFG		PEF		IDP		DKV	
		b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Models 1: Years spent teaching in gifted classes																	
of	HCAP school teachers[a]	-0.05	(0.07)	-0.10	(0.07)	-0.06	(0.08)	-0.06	(0.07)	-0.04	(0.08)	-0.01	(0.08)	-0.07	(0.07)	-0.05	(0.07)
	HCAP instructors	0.00	(0.06)	-0.04	(0.06)	0.08	(0.06)	0.02	(0.06)	-0.07	(0.06)	0.14*	(0.06)	0.17**	(0.06)	-0.01	(0.06)
Models 2: Years spent teaching in general classrooms																	
of	Elementary school teachers	0.19	(0.09)	0.14	(0.10)	0.14	(0.10)	0.16	(0.10)	0.12	(0.10)	0.07	(0.10)	-0.07	(0.10)	0.08	(0.10)
	HCAP school teachers[a]	0.09	(0.09)	0.12	(0.08)	0.00	(0.08)	0.18	(0.09)	0.10	(0.08)	-0.09	(0.08)	-0.04	(0.08)	-0.07	(0.09)
Model 3: Beliefs about the malleability of intelligence[b]																	
of	Elementary school teachers	-0.11	(0.09)	0.14	(0.09)	0.01	(0.09)	-0.21*	(0.08)	-0.43***	(0.07)	0.21*	(0.08)	0.28***	(0.08)	0.24***	(0.08)
	Student teachers	-0.12**	(0.04)	0.02	(0.04)	0.06	(0.04)	-0.12**	(0.04)	-0.35***	(0.04)	0.29***	(0.04)	0.30***	(0.04)	0.19***	(0.04)

Note. DHG = Domain-Specific vs. Holistic Giftedness; HH = Heterogeneity vs. Homogeneity of Gifted Children; IA = Importance of Achievement; II = Importance of Intelligence; MFG = Mutable vs. Fixed Giftedness; PEF = Interplay of Personal and Environmental Factors; IDP = Influence of Deliberate Practice; DKV = Different Key Variables for Children and Adults. Significance tests were set to an overall level of $\alpha = .05$ and adjusted with the Benjamini-Hochberg procedure (1995).

[a] HCAP school teachers had teaching experience in gifted classes and in general classrooms. Each of the two rows depicts results based on a model with only one kind of teaching experience. The statistical significance of the results did not change if both kinds of teaching experience were entered together into a model.

[b] High end = incremental theory of intelligence; low end = entity theory of intelligence. This scale was administered only to elementary school teachers and student teachers.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

that giftedness develops through an interplay of diverse factors, that giftedness can be achieved through deliberate practice, and that giftedness means different things for children and adults. Student teachers with more incremental beliefs about intelligence saw giftedness also as more domain-specific than holistic. No associations were found for beliefs about the malleability of intelligence and beliefs about whether giftedness has to be shown through excellent achievement or whether gifted students are a heterogeneous or homogeneous group. The associations between beliefs about the malleability of intelligence and beliefs about the nature of giftedness were not significantly different between the two groups (all *p*s > .05).

2.5 Discussion

In this study, we measured beliefs about the nature of giftedness in four groups of teachers: elementary school teachers, HCAP school teachers, HCAP instructors, and student teachers. We conceptualized these beliefs on the basis of key characteristics from Subotnik et al.'s (2011, 2012) definition of giftedness. We derived eight dimensions to measure different beliefs about the content and development of giftedness in alignment with aspects of giftedness that have been discussed in the scientific literature: *Domain-Specific vs. Holistic Giftedness, Heterogeneity vs. Homogeneity, Importance of Achievement, Importance of Intelligence, Mutable vs. Fixed Giftedness, Interplay of Personal and Environmental Factors, Deliberate Practice, and Different Key Variables for Children and Adults*. In the following, first, we discuss the main results for the structure of beliefs about the nature of giftedness, overall tendencies and mean-level differences in these beliefs, and their associations with beliefs about the malleability of intelligence. Second, we discuss the strengths and limitations of our study, and third, we draw a conclusion.

2.5.1 Beliefs about the nature of giftedness

2.5.1.1 The structure of beliefs about the nature of giftedness

We found empirical support for the proposed dimensions in the total sample (N = 1,235). The finding that all of the eight dimensions could be empirically supported suggests that—when asked directly—teachers basically have a structural representation of beliefs about the nature of giftedness that is in alignment with the aspects of scientific conceptions of giftedness that have been discussed in the literature. Furthermore, the low correlations between the dimensions indicate that the dimensions are theoretical distinct and rather independent. However, adaptations had to be made in order to establish the factor structure in the total sample. In particular, the problems in establishing the *Importance of Achievement* and *Importance of Intelligence* factors illustrated the need for additional detailed investigations. Both factors are highly relevant for conceptions of giftedness, and many conceptions can be separated by their standing on these two factors (see Sternberg & Davidson, 2005). The problem with these two dimensions was not that the items were intermingled; rather, the problems occurred within the dimensions. This might indicate that teachers have different or more differentiated conceptions concerning these two dimensions in relation to the approach that was applied.

We found only partial strong measurement invariance for the four groups of teachers. We followed Chen’s (2008) analysis for identifying reasons for a lack of measurement invariance as we explored patterns in our lack of invariance. First, we found an eight-factor structure within each of the four groups. This can be taken as an indicator that all four groups have similarly differentiated beliefs about the nature of giftedness in relation to the range of items that we measured. However, we found that some of items’ factor loadings but not their intercepts had to be freed to establish partial strong measurement invariance. Reasons might be that the content of the noninvariant items was understood differently by some of the groups or might be connected with different (strong) associations. Another possibility for the invariance might also be that the groups of teachers had different response sets.

2.5.1.2 Tendencies in beliefs about the nature of giftedness

We observed some general tendencies in teachers’ beliefs about the nature of giftedness. Most results were in line with modern conceptions of giftedness such as Subotnik et al.’s (2011, 2012), especially when considering that teachers were exclusively asked about gifted children and not adolescents or adults. In consensus with Subotnik et al., who stated that the key variable for giftedness during the early stages is the potential to achieve rather than actual achievement, teachers in our study tended to disagree with the statement that giftedness must be shown through superior achievement. Further, teachers tended to agree on average that intelligence is the most important characteristic of giftedness.

However, only the belief that giftedness can be developed through deliberate practice was rather clearly negated on average, whereas all other factor means were located closer to the midpoint of the scale. Furthermore, the variance was rather large and the correlations between the dimensions rather low. In conclusion, although we were able to observe some overall tendencies that teachers' beliefs were in line with modern conceptions of giftedness, the results also suggested that teachers had diverse beliefs about giftedness. It is an open question whether the reported diversity of beliefs is due to different profiles of beliefs that are or are not in line with different streams of conceptions, a confusion about the meaning of giftedness, or the belief that nearly everything is possible concerning the conception of giftedness.

2.5.1.3 Mean-level differences in beliefs about the nature of giftedness

In relation to differences in teachers' beliefs about the nature of giftedness, two results stand out. First, student teachers' beliefs were markedly different from the beliefs of the other groups. This result might be associated with differences in the amount of time spent teaching gifted students such as found in studies by Endepohls-Ulpe and Ruf (2005) and Schack and Starko (1990) or with differences in the amount of time spent teaching in general classrooms as found by Guskin et al. (1992). However, to understand whether and how these variables might be associated with beliefs about the nature of giftedness and the differences that we observed between the groups, there is a need to examine—according to Fives and Buehl's (2012) identification of dimensions on which beliefs can differ—whether the beliefs we measured are stable or dynamic, how these beliefs line up with knowledge about giftedness, and the positions of these beliefs within teachers' belief systems. Furthermore, different understandings of some of the items might explain (at least in part) the differences between the groups. The student teachers had the largest number of factor loadings that had to be freed in order to establish partial strong measurement invariance. Although all of the items referred to gifted children, it might be the case that student teachers thought more about older students because they were studying to become secondary school teachers. With this thought in mind, student teachers' beliefs might be more in line with the next stage of developmental conceptions (e.g., Ericsson, 1996; Subotnik et al., 2011, 2012) where the impact of intelligence is not as pronounced as in earlier stages but where training and practice in specialized domains are more relevant.

Second, none or only a few differences in beliefs about the nature of giftedness were found between and within the other groups (i.e., elementary school teachers, HCAP school teachers, and HCAP instructors). For elementary school teachers and HCAP school teachers, the results are in contrast with previous results (cf. Endepohls-Ulpe & Ruf, 2005; Schack & Starko, 1990) that showed differences that were related to the amount of time spent teaching gifted students. It might be the case that the amount of contact with gifted students in these special classes was not as influential as the amount of time spent teaching in general classrooms. The small associations between the amount of time HCAP instructors spent teaching gifted classes and their beliefs about the nature of giftedness might be due to the circumstance that the only place where they could obtain teaching experience with (gifted) students was in the HCAP.

2.5.1.4 Associations between beliefs about the nature of giftedness and beliefs about the malleability of intelligence

We found that beliefs about the malleability of intelligence were related to beliefs about the nature of giftedness. Respondents who viewed intelligence as more malleable and changeable also tended to have similar views of giftedness. Clearly, this finding can be aligned with the fact that intelligence is—with a few exceptions—a significant component of scientific giftedness models (e.g., Heller, Perleth, & Lim, 2005; Mönks & Katzko, 2005; Subotnik et al., 2011) and, moreover, is a characteristic of gifted students often mentioned by teachers (e.g., Endepohls-Ulpe & Ruf, 2005; Hany, 1997). Thus, the correlations indicate that student teachers and elementary school teachers tend to incorporate beliefs about intelligence into their belief systems about giftedness. Therefore, beliefs about the malleability of intelligence should be considered, for example, when trainings in the area of giftedness are established. In light of the diverse associations that have been found for beliefs about the malleability of intelligence with students' learning and teachers' professional positions (e.g., Jones et al., 2012; Rattan et al., 2012), beliefs about the malleability of giftedness should be investigated for possible associations with teachers' educational behavior toward gifted students.

2.5.2 Strengths and limitations

Our results and their interpretations should be considered in the light of several strengths and limitations of the current investigation. First, we investigated beliefs about giftedness in a large sample consisting of 1,235 participants from a German state. The participants belonged to four groups that are crucial for the (academic) development of gifted children, namely, elementary school teachers, enrichment program teachers and instructors, and student teachers. Our results appear to be relatively representative for these groups in this state. However, further studies are needed to investigate the generalizability of our results to other countries, educational programs, or school forms. Furthermore, the selected groups differed on variables that we could not control for. Therefore, the reasons underlying the differences in beliefs about the nature of giftedness remain to be elucidated.

Second, we based the construction of our questionnaire for measuring beliefs about the nature of giftedness on a current comprehensive conceptual framework—namely, the definition of giftedness proposed by Subotnik and her colleagues (2011, 2012). On the one hand, by doing so, we ensured that our questionnaire could be linked to a current scientific understanding of giftedness (one that has already taken into account many years of giftedness research). On the other hand, there are of course several alternative conceptions of giftedness that might have resulted in a different conceptualization. Future research might thus aim to broaden the questionnaire presented herein with regard to other conceptions of giftedness.

Third, the eight-dimensional structure of beliefs about giftedness was supported by a good fit to the data. We applied (multigroup) ESEM and EWC analyses to test for measurement invariance. This approach has many advantages as it combines several of the best features of more traditional approaches such as EFA and CFA. For example, it yields more accurate estimations of correlations than traditional confirmatory factor analysis if nontarget factor loadings exist (Marsh et al., 2014). However, the interpretation of the dimensions is more difficult, and the calculation of traditional indices (e.g., Cronbach's alpha) is contraindicated due to the allowance of cross-loadings. Furthermore, we only reached partial strong measurement invariance between the groups and the detection of (further) noninvariance might have been obscured due to the small group sizes of elementary school teachers and rather unequal group sizes (Chen, 2007; Meade, Johnson, & Braddy, 2008). Therefore, again, the reported differences and similarities between the groups might be (at least partly) due to structural differences between the groups.

Fourth, with the development of the questionnaire, we took a step toward the measurement of beliefs about the nature of giftedness in a comprehensive but also relatively differentiated manner. Of course, as this instrument is new, more research is needed to validate and optimize it. In particular, there is a need to revise the two dimensions *Importance of Achievement* and *Importance of Intelligence*. Furthermore, more items per dimension would strengthen the psychometric quality of the questionnaire, at least in terms of classical test theory.

Fifth, more research seems needed to systematically investigate the impact of the beliefs about the nature of giftedness on actual teacher behavior such as the identification (see, e.g., Rothenbusch et al., 2016) and promotion of gifted children. In a similar vein, to better understand the found similarities and differences in the beliefs, researchers might investigate the impact of constructs such as teachers' experience or level of training in the area of giftedness.

Last but not least, our use of a self-report questionnaire provided an economical measure for tapping into teachers' beliefs about the nature of giftedness. Although we aimed to measure relatively explicit parts of beliefs, it is possible that less direct methods such as the use of vignettes would have provided further insights into underlying beliefs. A comparison of results from direct and indirect examinations of beliefs about the nature of giftedness might therefore be conducted.

2.5.3 Conclusion

In conclusion, our study provides insights into teachers' beliefs about the nature of giftedness on the basis of theoretically derived dimensions that are close to characteristics of scientific conceptions of giftedness. Student teachers' beliefs were markedly different from those of the other three groups of teachers whose beliefs were on average rather similar to each other. However, teachers expressed overall a huge variety of beliefs. Our results yield new starting points for examining beliefs about the nature of giftedness. On a theoretical level, different conceptions of giftedness imply different views on whom is gifted, why someone is gifted, and how giftedness can be nurtured. The presented dimensions of beliefs can be used to adapt training programs in the area of giftedness and to investigate educational implications of teachers' different standings on debated issues in the conceptualization of giftedness.

Appendix

Excluded Items From the Questionnaire of Beliefs About the Nature of Giftedness

	Excluded items	Reasons for exclusion
Factor 1:	*Domain-Specific vs. Holistic Giftedness*	
DHG05R	A child's giftedness mostly refers to few areas.	target-factor loading < .30 formed an extra factor with DHG6R
DHG6R	Gifted children have superior abilities over peers only in certain areas.	target-factor loading < .30 formed an extra factor with DHG05R
Factor 2:	*Heterogeneity vs. Homogeneity of Gifted Children*	
HH5R	Gifted children mostly do not have more similarities than not gifted children.	target-factor loading < .30
Factor 3:	*Importance of Achievement*	
IA5R	Underachievement—an achievement level that is lower than a child's potential—can also happen to gifted children.	target-factor loading < .30
Factor 4:	*Importance of Intelligence*	
II3R	A model that contains further facets in addition to intelligence can explain the giftedness phenomenon best.	target-factor loading < .30 cross-loading ≥ .30 on PEF
II4	There are no other components than high cognitive abilities that represent giftedness in children.	target-factor loading < .30
II5R	Regardless of the height of a child's cognitive abilities, if further conditions are not met, the child is not gifted.	target-factor loading < .30
Factor 5:	*Mutable vs. Fixed Giftedness*	
MFG4R	Giftedness develops.	target-factor loading < .30 cross-loading ≥ .30 on IDP
Factor 6:	*Interplay of Personal and Environmental Factors (PEF)*	
PEF5R	Biological components are the basis; other factors have little influence on whether a child is gifted.	target-factor loading < .30
Factor 7:	*Influence of Deliberate Practice (IDP)*	
IDP5	Children can be strongly formed through their environment so that children's giftedness can be a result of their environment.	cross-loading ≥ .30 on PEF

Note. R (as in DHG05R) = reverse-scored.

3

Study 2:
Exploring Reference Group Effects on Teachers’ Nominations of Gifted Students

Rothenbusch, S., Zettler, I., Voss, T., Lösch, T., & Trautwein, U. (2016). Exploring reference group effects on teachers’ nominations of gifted students. *Journal of Educational Psychology, 108,* 883–897. http://doi.org/10.1037/edu0000085

Abstract

Teachers are often asked to nominate students for enrichment programs for gifted children, and studies have repeatedly indicated that students' intelligence is related to their likelihood of being nominated as gifted. However, it is unknown whether class-average levels of intelligence influence teachers' nominations as suggested by theory—and corresponding empirical results—concerning reference group effects. Herein, it was hypothesized that, when students' individual fluid and crystallized intelligence scores were similar, students from classes with higher average levels of intelligence would have a lower probability of being nominated for an enrichment program for gifted children than students from classes with lower average levels of intelligence. Furthermore, we investigated whether three teacher variables would influence the expected reference group effect, namely, experience with giftedness, beliefs about the changeability of intelligence, and the belief that giftedness is holistic or domain-specific. In a study comprising data from 105 teachers and 1,468 of their (German) third-grade students, we found support not only for a positive association between students' individual intelligence scores and the probability that students would be nominated as gifted, but also, more importantly, for the proposed reference group effect: When controlling for individual levels of intelligence, students' probability of being nominated was higher in classes with lower average levels of intelligence. In addition, the results showed that this reference group effect was stronger when teachers saw giftedness as holistic than as domain-specific. Also, depending on teachers' kinds of experience with giftedness, the reference group effect varied in size.

Keywords: giftedness, intelligence, reference group effects, teacher beliefs, teacher nominations

Exploring Reference Group Effects on Teachers' Nominations of Gifted Students

3.1 Introduction

One of the biggest challenges of enrichment programs for gifted students is determining how to select the "right" (i.e., gifted) participants (Heller, 2004; Worrell & Erwin, 2011). Teachers are often involved in the selection of students for such programs (Coleman & Gallagher, 1995; Deku, 2013; Freeman & Josepsson, 2002; McClain & Pfeiffer, 2012), and research has indicated that teachers' nominations of gifted students are substantially associated with students' intelligence, achievement, or related constructs (e.g., Endepohls-Ulpe & Ruf, 2005; Schack & Starko, 1990). This focus is in line with most conceptions of giftedness that involve high intelligence and achievement (see Sternberg & Davidson, 2005; Sternberg, Jarvin, & Grigorenko, 2011; Subotnik, Olszewski-Kubilius, & Worrell, 2011, 2012). At the same time, however, research has indicated that many highly intelligent students are overlooked by teachers and are thus not nominated to participate in programs for gifted and talented education (e.g., Gagné, 1994; Hunsaker, Finley, & Frank, 1997; Siegle, Moore, Mann, & Wilson, 2010).

Research investigating the problem that some highly intelligent students are not seen as gifted by their teachers has mostly considered individual-level student variables such as age, gender, or social (i.e., ethnic and socioeconomic) background and has provided many fruitful insights in this regard (e.g., Bianco, Harris, Garrison-Wade, & Leech, 2011; Callahan, 2005; Ford, 1998; Speirs Neumeister, Adams, Pierce, Cassady, & Dixon, 2007). Importantly, however, this individual-level perspective might be too narrow. McBee (2010), for instance, found that nomination rates for a gifted education program fluctuated greatly between schools. More generally, research on reference group effects has indicated not only that average levels of ability computed at either the class or the school level affect students' self-concepts (Marsh, 1984, 1987; Marsh et al., 2008), an effect that has been dubbed the *Big-Fish-Little-Pond Effect* (BFLPE; Marsh & Parker, 1984), but also that such frames of reference apply to teachers' perceptions of students: Students from schools with higher average ability levels have been found to receive lower teacher-assigned grades than equally able students from schools with lower average ability levels (Marsh, 1987; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006).

In the present study, we applied this reasoning about reference group effects to the area of primary school teachers' nominations of gifted third graders in Germany. That is, besides investigating the importance of individual-level student intelligence for teachers' nominations, we tested whether the probability that third graders with equal levels of intelligence would be nominated as gifted by teachers depends on the average level of intelligence in their classrooms. Specifically, we expected that students from classes with higher average levels of intelligence would have lower chances of being nominated than students with similar levels of intelligence who were in classes with lower average levels of intelligence. In addition, we also explored whether this reference group effect would be moderated by teachers' experience with giftedness, beliefs about the changeability of intelligence, and beliefs about giftedness.

3.1.1 Identifying gifted students for enrichment programs

Enrichment programs for the gifted are one of several strategies that aim to support the development of gifted students. Enrichment programs provide gifted students with additional instructions or learning content that differs from the regular school curriculum (Olszewski-Kubilius & Thomson, 2012). According to meta-analytic results, enrichment programs have a positive impact on gifted students' development (Kulik & Kulik, 1992; Vaughn, Feldhusen, & Asher, 1991). When enrichment programs are introduced, important decisions need to be made about how giftedness should be conceptualized and how children should be selected to participate in such interventions.

Giftedness has been conceptualized either as the potential to achieve at a high level or as exceptionally high performance in one or more domains (Sternberg & Davidson, 2005; Sternberg et al., 2011) and in relation to social reference groups (e.g., high-functioning individuals in one domain; Subotnik, Olszewski-Kubilius, & Worrell, 2011). Debates about the role of intelligence in giftedness have shaped how conceptions of giftedness have evolved (see, e.g., Kaufman & Sternberg, 2008): Traditional conceptions of giftedness (e.g., Hollingworth, 1942; Terman, 1925) focused on general intelligence and, in turn, expected gifted children to show high abilities or performance across a wide array of academic domains. Other conceptions have emphasized the multidimensionality of intelligence and have, accordingly, viewed students as gifted when they have shown exceptionally high intelligence or reasoning abilities in a specific domain (e.g., Brody & Stanley, 2005; Thurstone, 1938). Other conceptions (e.g., Renzulli, 2005b; Sternberg, 2005) have advocated that a high level of (general or domain-specific) intelligence is still important and often a prerequisite for giftedness. However, intelligence alone is not sufficient; characteristics such as creativity and motivation should be considered as well. Recent conceptions (e.g., Feldhusen, 2005; Subotnik et al., 2011, 2012) additionally recognize the developmental character of giftedness and see intelligence as especially relevant in the early stages of the development of giftedness.

In line with many conceptions of giftedness, common sources that are used to identify students for programs to promote the gifted consist of intelligence and (domain-specific) achievement tests (e.g., Coleman & Gallagher, 1995; Hoge & Cudmore, 1986) but also questionnaires or tests that assess additional student characteristics (e.g., creativity, motivation; e.g., Heller & Perleth, 2008). It is typically recommended to use multiple sources and criteria in the identification process (e.g., Frasier & Passow, 1994; Friedman-Nimz, 2009; VanTassel-Baska, Feng, & Evans, 2007). Furthermore, it is often recommended to include teachers in the process because they have detailed insights into the learning-related characteristics and behaviors of students (Heller & Perleth, 2008; Jarosewich, Pfeiffer, & Morris, 2002; Renzulli, 2005b). In the US, teachers often function as “gatekeepers” in the identification process. For instance, the *2012-2013 State of the States* report (National Association of Gifted Children, 2013) indicated that the majority of states used a multicriteria model in the identification process and that the identification process was often initiated, among other criteria, after a teacher nominated a student as potentially gifted. Note that other sources have reported an even stronger reliance on teacher nominations (e.g., Deku, 2013; Freeman & Josepsson, 2002). Given the (important) role of teachers in many processes that are used to identify gifted students, it is crucial to understand what might influence teachers’ nominations of gifted students.

3.1.2 Teachers’ nominations of gifted students

Many studies have investigated teachers’ nominations of gifted students (e.g., Endepohls-Ulpe & Ruf, 2005; Hany, 1993, 1997; Harradine, Coleman, & Winn, 2014; Hunsaker et al., 1997; Neber, 2004). But whereas some early studies (e.g., Gear, 1976; Pegnato & Birch, 1959) seemed to imply that teachers are not good at identifying gifted students, a review of 22 studies (Hoge & Cudmore, 1986) indicated that the empirical data did not justify such a view. More recently, McBee (2006) concluded that the accuracy of teacher nominations had yet to be determined due to the still deficient state of research in this area. His study suggested that teacher nominations appeared to be relatively high in quality in general, but he also found variability that depended on student characteristics.

Students' intelligence or characteristics that are related to cognitive abilities seem to guide teachers' nomination decisions (Endepohls-Ulpe & Ruf, 2005; Hany, 1997). But other constructs such as curiosity, creativity, and motivation have also been found to be important (Frasier & Passow, 1994; García-Ros, Talaya, & Pérez-González, 2012; Hany, 1997; Schack & Starko, 1990). At the same time, some variables that tend to bias teachers' nomination decisions have been identified. For instance, Bianco et al. (2011) indicated that boys had higher chances of being nominated than girls. Also, students from families with a higher socioeconomic status have often been found to be more likely to be identified as gifted than socioeconomically disadvantaged students (e.g., Ambrose, 2002; Elhoweris, Mutua, Alsheikh, & Holloway, 2005; Speirs Neumeister et al., 2007).

Besides, teachers' level of experience with giftedness seems to be related to their nomination decisions: According to Gear (1978), teachers who received training in giftedness were more effective than teachers without training in giftedness (i.e., from their classes, trained teachers identified more students as gifted who were also identified as gifted by an intelligence test), but the two groups were similarly efficient (i.e., from the students who were identified as gifted by teachers with or without training, the percentage of students who were also identified as gifted by an intelligence test was equal). Rubenzer and Twaite (1978) reported that teachers with training in the area of giftedness identified more students as gifted than teachers without training. Likewise, Bianco and Leech (2010) used vignettes to describe potentially gifted students who either had or did not have a disability. They found that teachers of the gifted referred more students in general for gifted education than regular school teachers or special education teachers did.

Teachers' beliefs about giftedness have also been investigated with regard to labeling students as gifted or not (e.g., Sternberg & Zhang, 1995), the education of gifted students (e.g., Bégin & Gagné, 1994, 1995), and gifted students' personality (e.g., Preckel & Baudson, 2013a). Herein, we focused on two kinds of beliefs, namely, (a) teachers' beliefs about the changeability of intelligence (given the importance of intelligence for most giftedness conceptions) and (b) teachers' belief that giftedness is holistic or domain-specific.

Beliefs about intelligence have often been investigated in terms of whether intelligence is believed to be malleable and changeable (i.e., incremental theory) or fixed and unchangeable (i.e., entity theory; Dweck, Chiu, & Hong, 1995). In general, holding an incremental theory of intelligence has been linked to focusing on learning and seeing challenges as opportunities to learn, whereas individuals who agree with an entity theory seem to place more of an emphasis on performing well (or not poorly) to validate their own abilities, tend to avoid challenges that might call their abilities into question, and are less interested in learning (Carr, Rattan, & Dweck, 2012). According to a study by Jonsson, Beach, Korp, and Erlandson (2012), teachers who believe that intelligence is fixed have been found to favor a g-factor theory of intelligence. García-Cepero and McCoach (2009) investigated whether an incremental or entity theory was associated with teachers' preferences for either a method that was based solely on intelligence tests or a method that used multiple criteria to identify gifted students, but they did not find any link. However, given that intelligence is one of the main characteristics of giftedness (see Sternberg & Davidson, 2005; Sternberg et al., 2011) and that beliefs about the changeability of intelligence are connected with a diverse number of variables, more research seems to be needed to investigate whether or not, and if so, how these beliefs are associated with teachers' nominations of gifted students.

In addition, we focused on the belief that giftedness is holistic or domain-specific. Some teachers might—as in rather traditional conceptions of giftedness that focus on general intelligence (e.g., Hollingworth, 1942; Terman, 1925)—see giftedness as holistic, resulting in a high level of performance in a wide array of domains. On the other hand, teachers might see giftedness as domain-specific. This view would be in accordance with more recent conceptions (e.g., Sternberg, 2005; Subotnik et al., 2011) that advocate that there are gifted students who show exceptional (potential for) performance in a wide range of domains but that most gifted students demonstrate high functioning in only one or a few specific domains (for more discussion on these differences, consult, e.g., Borland, 2005; Feldhusen, 1986; Kaufman & Sternberg, 2008; Renzulli, 2005b). Indeed, Brighton, Moon, Jarvis, and Hockett (2007) indicated that teachers were not sure whether students who demonstrated exceptional abilities in only one domain were as gifted as students who showed such abilities in multiple domains. However, to improve the identification of gifted minority students, Callahan (2005) advocated that teachers need to see that giftedness "is not a trait that demands that a child exhibit outstanding abilities in all areas" (p. 100). In light of the important differences between a holistic versus a domain-specific view of giftedness, we tested the relation between this variable and teacher nominations as well.

Whereas research on teachers' nominations of gifted students has primarily focused on the influence of individual student characteristics and has considered teacher variables such as teachers' experience with giftedness and different teachers' beliefs as well, the school and class contexts have been considered less often. However, the school and class contexts have theoretical and practical importance for the identification of gifted students (e.g., Lohman & Gambrell, 2012; McBee, 2010; Subotnik et al., 2011). For instance, research has indicated that teachers' judgments of students are influenced by reference groups (e.g., Baudson, Fischbach, & Preckel, 2014; Marsh, 1987). Also, several experts in the field (Lohman, 2005; Lohman & Gambrell, 2012; Olszewski-Kubilius & Clarenbach, 2012) have suggested that reference groups should be considered when identifying gifted students not only on a national level but also on local and subgroup levels. In contrast to an approach that relies on only national norms, gifted students who are socioeconomically disadvantaged or have a minority status have greater chances of being identified when reference groups are considered on a local or subgroup level. The idea is that a student's potential to achieve can be better assessed in relation to similar students who come from the same school where they had similar learning opportunities than in relation to all students in a state or country.

Herein, we aim to provide a more in-depth investigation of the role of the class context in teachers' nominations of gifted primary school students for an enrichment program in Germany. Specifically, given the importance of cognitive capabilities for giftedness, we focus on the role of the average level of intelligence in a classroom.

3.1.3 Reference group effects

Reference group effects refer to the influences that groups can have on the perceptions or evaluations of specific individual variables. Reference group effects have been investigated in the field of education in general, though rather rarely with regard to teachers' choices of which students to nominate for enrichment programs for gifted students. The arguably most well-established reference group effect is the BFLPE (Marsh & Parker, 1984), which describes a negative reference group effect of class-average or school-average abilities in a particular domain on students' self-concept in the same domain. Furthermore, reference groups have been found to influence teacher judgments.

In this regard, Baudson et al. (2014) investigated the influence of class-average levels of cognitive abilities on teachers' judgments of students' cognitive abilities. Controlling for students' individual ability levels, they found that teachers' judgments were negatively influenced by class-average levels of cognitive ability. Further, studies by Marsh (1987) and Trautwein et al. (2006) showed that students from classes with higher average achievement levels received lower teacher-assigned grades than equally able students from classes with lower average achievement levels. Similarly, Südkamp and Möller (2009) used a computer program to simulate a classroom and asked student teachers to rate fictitious students by giving them grades on a test and by estimating the numbers of questions they answered correctly. This study indicated that the assigned grades were affected by the simulated classmates: Students with identical achievement levels received lower grades in simulated classes with high average achievement levels than in classes with lower average achievement levels. However, only the grades but not the estimations of the numbers of correct answers were influenced by the reference group.

More evidence for the effects of reference groups on teachers' judgments has been accumulated for the tracked school systems in Germany and Switzerland. In these countries, there are different school tracks that cater to students with different achievement levels and that lead to different school leaving certificates. Typically, at the end of primary school, primary school teachers recommend students to a certain secondary school track (i.e., they assign students to different schools on the basis of the students' achievement levels). The studies that have been conducted so far have indicated that these so-called teacher transition recommendations are influenced by reference group effects (Baeriswyl, Wandeler, & Trautwein, 2011; Milek, Lüdtke, Trautwein, Maaz, & Stubbe, 2009; Tiedemann & Billmann-Mahecha, 2007; Wagner, Helmke, & Schrader, 2009). Equally able students from different classes received different recommendations for secondary schools in relation to their class' average levels of achievement such that students from classes with higher average achievement levels had higher probabilities of receiving recommendations for lower secondary school tracks than equally able students from classes with lower average achievement levels.

Finally, there is some evidence that reference groups might also affect teachers' nominations of giftedness. McBee (2010) investigated the nomination probabilities of primary school students who could be nominated in different ways—one of them was via teacher nominations. He found that different compositions of students in a school (e.g., a high percentage of Asian students or a high school-average achievement level) had a positive effect on students' probability of being nominated as gifted. Yet, in this study, he could not control for students' individual levels of achievement because he had information only about whether a student had been held back a grade but not more suitable variables. Building on this stream of research, we herein focused on teachers' nominations of students for an enrichment program for gifted students and expected that reference group effects would play a major role in students' probability of being nominated.

3.1.4 Potential moderators of this reference group effect

Some studies have explored whether reference group effects are also moderated by other variables. In general, only a relatively small number of studies have indicated such moderator effects. Most of these moderators have been individual-level variables (e.g., Jonkmann, Becker, Marsh, Lüdtke, & Trautwein, 2012; Seaton, Marsh, & Craven, 2010), whereas studies that have tested for moderator variables at the class or school levels have rarely found any moderating effects (Lüdtke, Köller, Marsh, & Trautwein, 2005). Herein, we investigated whether potential effects of reference groups on teachers' nominations might be moderated further. Specifically, we looked at three teacher variables as potential moderators, namely, teachers' experience with giftedness, teachers' beliefs about the changeability of intelligence, and teachers' beliefs that giftedness is holistic or domain-specific. We focused on these variables because they have been found to be important in the context of giftedness. However, these variables have not yet been investigated in terms of their potential influence on the association between class-average intelligence and students' probability of being nominated for an enrichment program.

3.2 Research Questions

Herein, we predominantly investigated whether reference groups affect teachers' nominations of gifted German primary school children. Furthermore, we were interested in the role of the three teacher-related variables in this context. In our study, teachers had very little information about their nomination task, and therefore, they had to rely on their beliefs about giftedness. We concentrated on students' probability of being nominated for an enrichment program. The accuracy of teachers' nominations was not assessed because they had no information about exact criteria.

In particular, we analyzed whether students' individual levels of intelligence and class-average levels of intelligence were associated with teachers' nominations of German primary school children. In line with general definitions of giftedness and previous research that has indicated that intelligence is an important indicator that teachers use to identify gifted students (e.g., Hany, 1997), we first hypothesized that students' fluid and crystallized intelligence would be positively associated with their probability of being nominated.

Second, in line with research on reference group effects in general and teacher-assigned grades (e.g., Südkamp & Möller, 2009) and transition recommendations in tracked school systems (e.g., Trautwein & Baeriswyl, 2007) in particular, we hypothesized that teachers' nominations would also be affected by the class-average levels of fluid and crystallized intelligence. More precisely, we hypothesized that students' probability of being nominated as gifted would be smaller when they were in classes with higher average levels of intelligence than equally intelligent students from classes with lower average intelligence levels.

Third, to investigate possible reference group effects of class-average intelligence in greater detail, we also considered three giftedness-related teacher variables. Specifically, we tested on a rather exploratory level whether (a) teachers' experience with giftedness, (b) their beliefs about the changeability of intelligence, and (c) their belief that giftedness is holistic or domain-specific would be associated with the size of the reference group effect or with students' probability of being nominated. We expected that all three variables would be significantly associated with the size of the reference group effect and with the nomination probability because (a) experience with giftedness has been previously found to be connected with teachers' nominations (e.g., Bianco & Leech, 2010), (b) teachers' belief that intelligence is fixed has been shown to be connected with a preference for a g-factor theory of intelligence (Jonsson et al., 2012), and (c) teachers who see giftedness as holistic in contrast to domain-specific might rely more strongly on students' intelligence as an indicator of giftedness and, therefore, might be more influenced by possible reference group effects of class-average intelligence.

3.3 Method

3.3.1 Participants

In 2010, the Hector Children's Academy Program (HCAP) was established in the German state of Baden-Württemberg. Today, 60 academies, typically located at one or more primary schools, participate in the HCAP. The goal of the HCAP is to offer enrichment classes for the 10% most gifted primary school children in Baden-Württemberg. The HCAP provides classes in all academic domains but has a more pronounced focus on the STEM (Science, Technology, Engineering, and Mathematics) subjects. The HCAP classes are voluntary and take place outside the regular school curriculum, either after school or on weekends. All primary schools in Baden-Württemberg are allowed to nominate children for the HCAP.

In the present study, we used a subsample from a larger HCAP effectiveness investigation that took place in the 2012-2013 school year. More specifically, in the larger investigation, primary schools all over Baden-Württemberg were randomly selected to participate from a pool of schools that had nominated students for the HCAP in previous school years. This strategy was implemented to avoid having a large percentage of schools in the sample from which no children ended up being nominated for the HCAP at all.

At each of the selected schools, we aimed to run assessments in two classes of third graders. We chose to focus on third graders because first and second graders are not yet able to read or understand a wide range of written test and questionnaire items by themselves (thus limiting us in the selection of tests and questionnaires for group-based assessments) and because fourth graders were not appropriate for the aims of the larger investigation. When there were more than two classes of third graders in a school, we again drew a random sample of two classes from all classes at a particular school. In turn, in schools with one or two classes of third graders only, we aimed to use all of the available classes per school. Finally, within each class, we aimed to obtain a full sample of all students in the class.

The data used in our study were collected prior to teachers' nominations of gifted students. The subsample used herein consisted of teachers (and their classes) who reported that they nominated students in the participating classes for the HCAP. Thus, we excluded data from teachers (and their classes) who reported that they were not in charge of nominating students from their class for the HCAP (e.g., when a class got a new teacher over summer vacation, the nomination was typically done by the former class teacher). The teachers in our sample (N = 105; 80% female; mean age = 44 years, SD = 11) came from 75 schools. Of these 75 schools, 45 schools had one participating teacher (class) each, and 30 schools had two participating teachers (classes) each. Concerning student data, we obtained information from 1,468 third graders (51% female; mean age = 8 years, SD = 1). On average, 13.98 (SD = 4.53) students participated per class (the participation quota per class was M = 71%, SD = 20%, Min = 13%, Max = 100%). Finally, for 1,116 (76%) of the 1,468 third graders, we had information on their socioeconomic status from their parents.

3.3.2 Procedure

The study was voluntary for all participants (i.e., students, parents, and teachers). Also, participants' privacy was protected via several steps, most importantly by using a professional provider to handle the assessments. Whereas our research team received completely anonymized raw data files only, the professional provider did not know the content of the tests and questionnaires that were administered.

After obtaining parental agreement, students worked on the tests and questionnaires during normal school hours. Students' age and gender were collected from lists of participating students provided by the schools (i.e., official data); this information was further checked against information provided by both the parents and the students themselves. Teachers were asked to complete a questionnaire parallel to the testing of the students. Importantly, the timing of the administration of the questionnaire was completely unrelated to that of the nomination process for the HCAP (which took place later). Parents' occupation (to obtain information on students' socioeconomic background) was assessed via a parent questionnaire. Students were given this parent questionnaire at the testing sessions and asked to pass it on to their parents. Parents could then fill out the questionnaires at home and send them back anonymously and independently via a stamped and addressed return envelope (to the professional provider).

Table 3.1

Descriptive Statistics for Primary School Students Who Were Nominated or Not Nominated

Students	*N*	Age in years	Female	HISEI	Gf	Gc
	%	*M* (*SD*)	%	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Not nominated	77.60	8.67 (0.48)	51.58	54.13 (19.24)	-0.15 (0.96)	-0.12 (0.98)
Nominated	19.61	8.52 (0.39)	45.83	63.88 (17.31)	0.56 (0.96)	0.49 (0.95)

Note. N = 1,468 third graders; 2.79% had missing data on the variable *Nomination*. HISEI = Highest International Socio-Economic Index of Occupational Status; Gf = fluid intelligence; Gc = crystallized intelligence. Gf and Gc were z-standardized.

3.3.3 Measures

3.3.3.1 Student variables

We assessed several student variables that were pertinent to the current investigation: students' age, gender, socioeconomic status (as control variables), fluid and crystallized intelligence, and whether students were nominated by their teachers to participate in the HCAP (see Table 3.1 for additional information about the student characteristics).

3.3.3.2 Nomination

Teachers were asked to nominate students for the HCAP. They did not receive specific instructions or rating scales on which to base their decisions but were informed about the goal of the HCAP (i.e., providing enrichment classes for the 10% most gifted primary school children in Baden-Württemberg, Germany). The nomination was a global and undifferentiated judgment of students that was executed by simply registering students at an academy (either electronically or in written form). Teachers did not receive any feedback about students' test scores, nor did they receive any direct information from any academy about whether a nominated student was finally accepted for participation or not. The teachers nominated 288 of the 1,468 students (19.62%), and each teacher nominated 2.74 students (SD = 2.07) on average. Note that nominated students did not undergo further assessment. In the year of the investigation, virtually all nominated children finally attended a course at an academy (the correlation between "nomination" and "course attendance" in our sample was $r = .91$, $p < .05$).

3.3.3.3. Socioeconomic status

Families' socioeconomic status was measured with the *Highest International Socio-Economic Index of Occupational Status* (HISEI; M = 56.56, SD = 19.19). For each student, this is the score of the parent with the higher ISEI-08 (Ganzeboom, 2010). The ISEI-08 scores were calculated from the *International Standard Classification of Occupation 2008* (ISCO-08) scores. The scores were originally coded in an older format (ISCO-88), which was converted into the ISCO-08 format with SPSS syntax provided by Ganzeboom and Treiman (2011). ISEI-08 values contained information about parents' occupations. Scores could range from 10 for occupations such as "subsistence farmers, fishers, hunters, and gatherers" (lowest socioeconomic status) to 89 for occupations such as "medical doctors" (highest socioeconomic status).

3.3.3.4 Intelligence

Intelligence was measured with an adaptation of the German intelligence test *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe* (*Berlin Test of Fluid and Crystallized Intelligence for Grades 8–10,* BEFKI 8-10; Wilhelm, Schroeders, & Schipolowski, 2014) for Grades 3 and 4 (BEFKI 3-4). There are two parallel versions with the same types of tasks, each consisting of three subtests (34 items in total) that measure the verbal, numeric, and figural parts of fluid intelligence ($M_{Version\ A}$ = 16.12, SD = 5.34; $M_{Version\ B}$ = 15.76, SD = 5.22) and 42 items for assessing declarative knowledge (i.e., crystallized intelligence; $M_{Version\ A}$ = 19.15, SD = 5.69; $M_{Version\ B}$ = 18.52, SD = 5.72). The reliabilities of the different versions of the BEFKI 3-4 were good in terms of Cronbach's α for this sample (fluid intelligence: $\alpha_{\text{Version A}}$ = .77, $\alpha_{\text{Version B}}$ = .76; crystallized intelligence: $\alpha_{\text{Version A}}$ = $\alpha_{\text{Version B}}$ = .74). Each child was given one version of the BEFKI 3-4. The versions were randomized across classes. The fluid and crystallized intelligence scores for each version were then z-standardized to place them on the same metric and combined into one fluid intelligence variable and one crystallized intelligence variable.

3.3.3.5 Class-average intelligence levels

We aggregated the students' z-standardized fluid and crystallized intelligence scores at the class level (fluid intelligence: M = 0.00, SD = 0.41; crystallized intelligence: M = -0.01, SD = 0.41). Thus, the score for one class on one of the variables for class-level intelligence was computed as the average of the single intelligence scores of each student in this class.

3.3.3.6 Teacher variables

The three teacher characteristics were assessed via a teacher questionnaire.

Giftedness experience. Teachers' experience with giftedness was measured with one item that could be answered yes or no: "Have you already dealt with giftedness in the context of your training or occupation as a teacher?" Of the participating teachers, 38.10% answered yes to this question. Teachers who answered "yes" were then asked to specify their experience using five given response options plus "other." The results were as follows for teachers who had experience with giftedness (multiple answers were possible): 22.5% of the teachers had attended lectures or courses about giftedness while studying at their university, 25% of the teachers had received on-the-job training, 7.5% of the teachers had received off-the-job training, 17.5% of the teachers offered education for gifted students, and 50% of the teachers read literature about giftedness. Thirty percent of the teachers marked "other" and usually specified that they had gifted students in their classes.

Beliefs about the changeability of intelligence. Beliefs about intelligence were measured via a translated version of the *Theories of Intelligence Scale – Self Form for Adults* (Dweck, 1999) consisting of eight items (M = 2.88, SD = 0.83) with a 6-point Likert-type scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). This questionnaire measures the degree to which respondents believe that intelligence is a fixed and immutable trait (entity theory) or that it is changeable and malleable (incremental theory). Four of the items describe agreement with entity theory, and four of the items describe agreement with incremental theory. The responses to the items referring to entity theory are recoded when creating an individual's test score. Hence, lower values indicate an entity theory of intelligence, and higher values indicate an incremental theory of intelligence. A sample item is "To be honest, people can't really change how intelligent they are" (recoded). The Cronbach's alpha reliability of the scale in this sample was good (α = .90). We also calculated the short form of this scale using the four entity theory items only (see Dweck, 1999). However, using this subset did not change the pattern of our findings; hence, we report only the results for the full scale.

Belief that giftedness is holistic or domain-specific. To investigate teachers' beliefs about whether giftedness is holistic or domain-specific, we used four items (with a 4-point Likert-type scale ranging from 1 = *disagree* to 4 = *agree*) measuring two opposing beliefs about giftedness: The low end represents the belief that giftedness is domain-related, and the high end represents the belief that giftedness is a holistic characteristic that simultaneously manifests itself across multiple domains. The items are "Gifted children can be average or below average in some areas in comparison with other children of the same age" (recoded), "Theoretically, a gifted child has a high potential to achieve in nearly all academic areas," "Gifted children are good in most school subjects," and "Gifted children are at the top of the class in almost all ability areas." The reliability of the scale was $\alpha = .76$ in this sample ($M = 2.02$, $SD = 0.56$). This scale is part of a newly developed questionnaire measuring different aspects of beliefs about giftedness and has been shown to have satisfactory to good measurement properties overall (Rothenbusch, Zettler, Lösch, & Voss, 2016).

3.3.4 Analyses

We ran analyses with the Mplus 7.11 software package (Muthén & Muthén, 1998-2012). Before performing the analyses, we standardized ($M = 0$, $SD = 1$) all continuous variables measured on students, classes, and teachers but retained the original metric of dichotomous variables. For missing data, which ranged from 1.90% for teachers' beliefs about intelligence to 29.86% for students' HISEI scores, we applied the full information maximum likelihood (FIML) algorithm. The FIML algorithm uses all available information when computing matrices for parameter estimates (Enders, 2010), providing acceptable to good estimations of regression coefficients and standard errors for the percentages of missing data in our study (Buhi, Goodson, & Neilands, 2008; Enders, 2001; Johnson & Young, 2011; Schlomer, Bauman, & Card, 2010). Because multiple significance testing increases the probability of falsely rejecting the null hypothesis (i.e., inflated Type I error rate), we applied the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) to control the false discovery rate.

3.3.4.1 Hierarchical data structure

In general, students in the same class are more likely to be similar to each other than to students in different classes (i.e., students are nested in school classes). If such a nested structure is ignored, small *Intraclass Correlations* (ICC) of .01 or .05 could already lead to a higher alpha error and smaller confidence intervals in conventional OLS regressions (Cohen, Cohen, West, & Aiken, 2003; Snijders & Bosker, 2012). In the present study, the ICCs varied from .01 for students' gender to .10 for fluid and crystallized intelligence. The *type = complex* and *type = twolevel* procedures in Mplus 7.11 are equally adequate for dealing with nested data. However, due to the lack of detailed documentation on the algorithms that Mplus 7.11 applies in multilevel logistic regression analyses, we used the well-established type = complex procedure to adjust the standard errors of the regression coefficients (for more information, see Muthén & Satorra, 1995).

3.3.4.2 Stepwise logistic regression

Logistic modeling with the MLR estimator was chosen because the outcome variable was binary (nominated as gifted vs. not nominated as gifted). In logistic regression analyses, the residual variance is fixed. Therefore, the total variance differs between models according to the amount of explained variance. In conclusion, only the direction but not the strength of model coefficients are comparable between (even nested) models when additional residual variables are not used to make adjustments (for more information, see Karlson, Holm, & Breen, 2012). However, we did not apply the model adjustments for model comparability suggested by Karlson et al. (2012) because we were not interested in differences between the coefficients of different models. Using the logit coefficients, we calculated odds ratios (OR). The odds are the probability of being nominated divided by the probability of not being nominated. The OR is the effect size of a variable's effect in a logistic regression. It indicates the increase in the odds of a student being nominated with a one standard deviation increase in one of the predictor variables in comparison with the odds of a student being nominated without the one standard deviation increase in the predictor variable (controlling for all other variables).

We used a stepwise sequential modeling approach with increasing complexity. The models were calculated separately for students' fluid and crystallized intelligence in order to avoid multicollinearity. To analyze the impact of students' intelligence on students' probability of being nominated, we first entered students' individual intelligence scores (Models 1a and 1b), and in a second step, we included additional student variables as control variables (Models 2a and 2b). To investigate reference group effects, in the third step, we then entered the class-average intelligence level (Models 3a and 3b). Therein, a

Table 3.2

Correlations Between Variables

Students	Nom	Age	Gender	HISEI	Gf	Gc	C-Gf
Nomination status (0 = not nominated, 1 = nominated; Nom)	—						
Age	$-.21^{***}$	—					
Gender (0 = male, 1 = female)	-.07	$-.08^{**}$	—				
HISEI	$.30^{***}$	$-.21^{***}$	-.01	—			
Fluid intelligence (Gf)	$.39^{***}$	$-.13^{***}$	-.01	$.27^{***}$	—		
Crystallized intelligence (Gc)	$.33^{***}$	-.06	-.04	$.22^{***}$	$.49^{***}$	—	
Class-average fluid intelligence (C-Gf)	-.01	$-.10^{***}$	.03	$.18^{***}$	$.40^{***}$	$.26^{***}$	—
Class-average crystallized intelligence	.04	$-.08^{**}$	.00	$.15^{***}$	$.25^{***}$	$.41^{***}$	$.64^{***}$
Teachers	**Nom**	**GE**	**BI**				
Giftedness experience (0 = no, 1 = yes; GE)	-.08	—					
Beliefs about the changeability of intelligence (BI)	-.05	$.23^{*}$	—				
Belief that giftedness is holistic or domain-specific	.07	.02	-.22				

Note. $N_{teachers}$ = 105; $N_{students}$ = 1,468. Significance tests were set to an overall level of α = .05 and adjusted with the Benjamini-Hochberg procedure (1995).
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

significant negative logit coefficient would indicate a negative reference group effect in line with our theorizing. In the fourth step, we used separate models for fluid and crystallized intelligence to test each teacher variable (i.e., two models were calculated for each teacher variable) for whether the teacher variables had direct effects on students' probability of being nominated while controlling for individual students' age, gender, and HISEI (and intelligence). In the fifth step, we analyzed the interaction terms (again, using separate analyses as in the fourth step). They were defined on the basis of the manifest teacher variables and the manifest variables for class-average intelligence.

3.4 Results

3.4.1 Association between students' individual intelligence and teachers' nominations

The correlations between the variables used in the present study can be found in Table 3.2. For our first hypothesis, we examined the association between students' intelligence and their probability of being nominated (see Table 3.3). In Models 1a and 1b, we entered students' fluid and crystallized intelligence scores separately as predictors. Each showed a positive association with students' probability of being nominated (b = 0.76, OR = 2.14, $p < .05$; b = 0.65, OR = 1.92, $p < .05$). Next, we added the control variables (Models 2a and 2b), namely, age, gender, and HISEI. Students' fluid and

Table 3.3

Students' Intelligence and Average Intelligence at the Class Level on Teachers' Nominations for an Enrichment Program

	Model 1a		Model 1b		Model 2a		Model 2b		Model 3a		Model 3b	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Intercept	-1.53***	.10	-1.50***	.10	-1.43***	.11	-1.43***	.12	-1.48***	.12	-1.45***	.12
Student variables												
Fluid intelligence	.76***	.08			.67***	.08			.86***	.09		
Crystallized intelligence			.65***	.09			.56***	.09			.68***	.09
Control variables												
Age					-.25**	.08	-.28***	.08	-.27**	.08	-.30***	.08
Gender (0 = male, 1 = female)					-.33*	.15	-.27	.14	-.33*	.15	-.26	.14
Socioeconomic status					.40***	.09	.42***	.09	.43***	.09	.45***	.09
Classroom variables												
Class-average fluid intelligence									-.47***	.10		
Class-average crystallized intelligence											-.29**	.10
Pseudo-R^2	.15***	.03	.12***	.03	.22***	.03	.19***	.04	.26***	.03	.21***	.04

Note. $N_{teachers}$ = 105; $N_{students}$ = 1,468. Significance tests were set to an overall level of α = .05 and adjusted with the Benjamini-Hochberg procedure (1995).
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

crystallized intelligence remained significant positive predictors of students' probability of being nominated (b = 0.67, OR = 1.96, $p < .05$; b = 0.56, OR = 1.76, $p < .05$). In addition, younger students had higher chances of being nominated (b = -0.25, OR = 0.78, $p < .05$; b = -0.28, OR = 0.75, $p < .05$), and being a girl decreased students' probability of being nominated in Model 2a, which included fluid intelligence (b = -0.33, OR = 0.72, $p < .05$), but not in Model 2b, which included crystallized intelligence. Students' socioeconomic status was positively associated with students' probability of being nominated (b = 0.40, OR = 1.49, $p < .05$; b = 0.42, OR = 1.53, $p < .05$). Overall, the results supported our first hypothesis that students with higher intelligence scores had a higher probability of being nominated—even when controlling for students' age, gender, and socioeconomic status, all of which also affected teachers' nominations.

3.4.2 Negative effects of the reference group on teacher nominations

Next, we added class-average fluid intelligence (Model 3a) and class-average crystallized intelligence (Model 3b) to the student predictors (see also Table 3.3). Controlling for the other student variables, both class-average fluid intelligence (b = -0.47, OR = 0.63, $p < .05$) and class-average crystallized intelligence (b = -0.29, OR = 0.75, $p < .05$) had a negative impact on students' probability of being nominated. These results indicated a negative effect of the reference group in terms of the class-average levels of fluid and crystallized intelligence on teachers' nominations. If a student came from a class with a higher average level of fluid or crystallized intelligence, her or his probability of being nominated was lower than the probability of a student with equal characteristics who came from a class with a lower average level of fluid or crystallized intelligence. The results thus supported our second hypothesis.

3.4.2 Effects of teacher variables on teacher nominations and reference group effects

Finally, we examined on an exploratory level the associations between the following teacher variables and students' probability of being nominated and, more specifically, their effects on the relation between class-average intelligence and students' probability of being nominated: (a) experience with giftedness as a general yes-or-no variable that was further subdivided into six kinds of experience, (b) beliefs about the changeability of intelligence, and (c) the belief that giftedness is holistic or domain-specific.

To analyze the direct effects of the teacher variables on students' probability of being nominated, we extended Models 3a and 3b by using the teacher variables as additional predictor variables. Each variable was entered separately into the regression equation. Furthermore, for each teacher variable, we ran two separate models: one for fluid and one for crystallized intelligence. Thus, a total of 18 models were specified to separately investigate the effects of the seven variables that measured experience with giftedness, the one variable that measured beliefs about the changeability of intelligence, and the one variable on the belief that giftedness is holistic or domain-specific.

In comparison with Model 3a and Model 3b, the direction and statistical significance of the regression coefficients for students' intelligence, class-average intelligence, age, gender, and HISEI remained exactly the same. Only one teacher variable was significantly related to students' probability of being nominated in terms of a main effect: In the analyses with crystallized intelligence, students had a lower probability of being nominated if they had teachers with more experience with giftedness (b = -0.43, OR = 0.65, p < .05). None of the other regression coefficients for the teacher variables were statistically significant.

In the next and final step, we extended the models that tested the direct effects of the teacher variables to examine whether their interactions with class-average intelligence were related to students' probability of being nominated. Again, we ran two separate models for each teacher variable and its interaction term: one for fluid and one for crystallized intelligence. Interaction terms were calculated using the manifest teacher variables and the manifest variables for class-average intelligence. A total of 18 models were calculated to separately investigate the seven interaction terms for the different experience with giftedness variables, one interaction term for beliefs about the changeability of intelligence, and one interaction term for the belief that giftedness is holistic or domain-specific.

Again, the direction and statistical significance of the regression coefficients for students' intelligence, class-average intelligence, age, gender, and HISEI were the same as in Model 3a and Model 3b. Furthermore, the addition of the interaction effects did not change the direction or level of significance of the main effects of the teacher variables. In the model with crystallized intelligence, students' probability of being nominated was lower if their teachers had experience with giftedness (b = -0.43, OR = 0.65, $p < .05$). Whether teachers had or did not have experience with giftedness was not related to the negative link between students' probability of being nominated and both class-average fluid intelligence and class-average crystallized intelligence. However, four interaction effects emerged for the different kinds of experience with giftedness (see Figure 3.1): The negative reference group effect of class-average fluid and crystallized intelligence on students' probability of being nominated was smaller if teachers had attended lectures or courses about giftedness while they were studying at their university (model for fluid intelligence: $b_{moderator}$ = -0.28, OR = 0.76; $p > .05$; $b_{interaction}$ = 0.73, OR = 2.07; $p < .05$; model for crystallized intelligence: $b_{moderator}$ = -0.41, OR = 0.67; $p > .05$; $b_{interaction}$ = 0.92, OR = 2.51; $p < .05$). Further, having read literature about giftedness also reduced the negative effect of class-average crystallize intelligence ($b_{moderator}$ = -0.42, OR = 0.66; $p > .05$; $b_{interaction}$ = 0.52, OR = 1.68; $p < .05$). By contrast, the negative effect of class-average fluid intelligence on students' probability of being nominated was even more negative if teachers had received off-the-job training in giftedness ($b_{moderator}$ = 0.33, OR = 1.39, $p > .05$; $b_{interaction}$ = -0.69, OR = 0.50; $p < .05$). In the model with crystallized intelligence, the interaction term was not significant, but students' probability of being nominated was higher if their teachers had received off-the-job training (b = 0.98, OR = 2.65, $p < .05$).

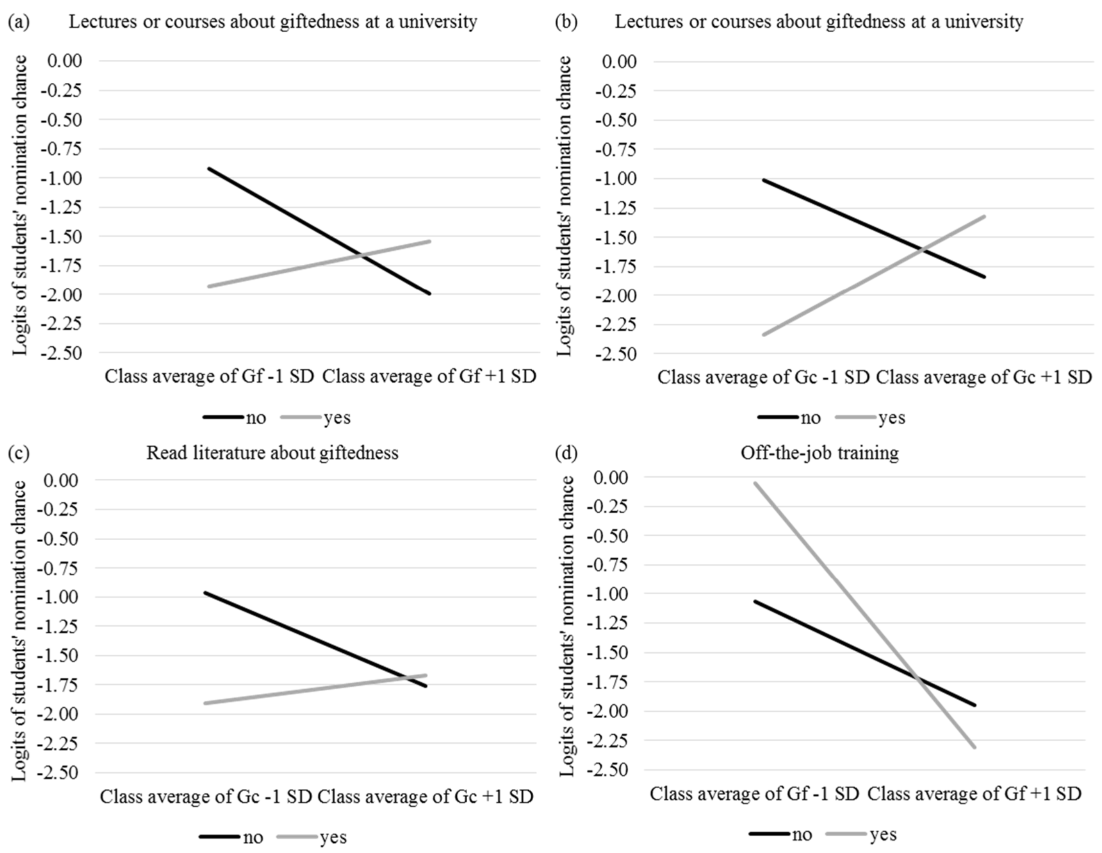

Figure 3.1. Influence of different kinds of experience with giftedness on the negative impact of class-average intelligence on students' nomination probabilities: (a) interaction between "Lectures or courses about giftedness at a university" and "Class-average fluid intelligence," (b) interaction between "Lectures or courses about giftedness at a university" and "Class-average crystallized intelligence," (c) interaction between "Reading literature about giftedness" and "Class-average crystallized intelligence," and (d) interaction between "Off-the-job training" and "Class-average fluid intelligence." Class-average Gf = Class-average fluid intelligence; Class-average Gc = Class-average crystallized intelligence. No = teachers had no experience with the depicted kind of giftedness. Yes = teachers had experience with the depicted kind of giftedness. Logits have a sigmoidal association with probabilities. Logits of -2.50 to 0.00 correspond with probabilities of 7.59% to 50%.

The direct effect of teachers' beliefs about the changeability of intelligence on students' probability of being nominated and its interactions with class-average fluid and class-average crystallized intelligence were not statistically significant. By contrast, teachers' belief that giftedness is holistic or domain-specific significantly moderated the associations with both the class-average level of fluid intelligence ($b_{moderator}$ = 0.20, OR = 1.22, p > .05; $b_{interaction}$ = -0.26, OR = 0.77, p < .05) and the class-average level of crystallized intelligence ($b_{moderator}$ = 0.24, OR = 1.27, p < .05; $b_{interaction}$ = -0.31, OR = 0.73, p < .05). Both interactions are displayed in Figure 3.2. In both cases, the negative influence of class-average intelligence on students' probability of being nominated decreased the more teachers tended to see giftedness as domain-specific.

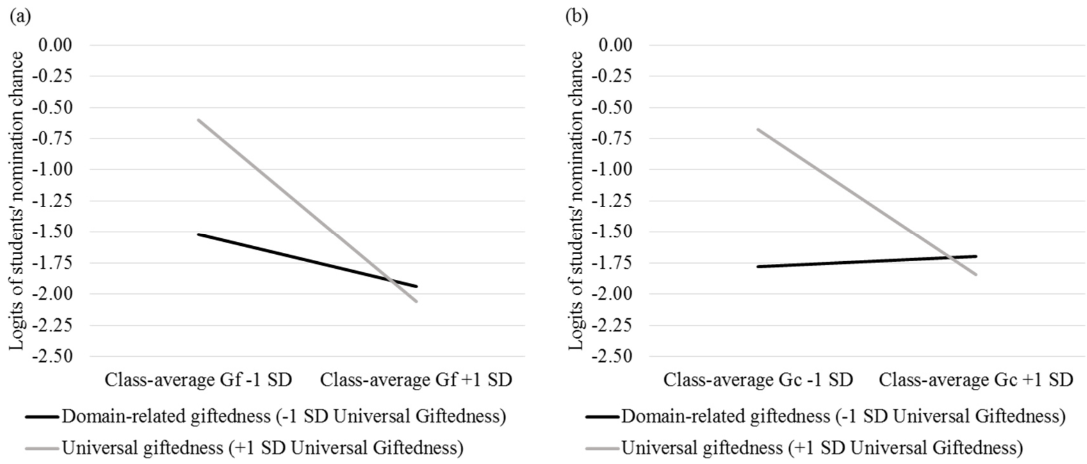

Figure 3.2. Influence of teachers' belief that giftedness is holistic or domain-specific on the negative impact of (a) class-average fluid intelligence on students' probability of being nominated and (b) class-average crystallized intelligence on students' probability of being nominated. Class-average Gf = Class-average fluid intelligence; Class-average Gc = Class-average crystallized intelligence. Logits have a sigmoidal association with probabilities. Logits of -2.50 to 0.00 correspond with probabilities of 7.59% to 50%.

Across all models, students' intelligence and HISEI were positively associated, and the class-average level of intelligence, students' age, and gender were negatively associated with students' probability of being nominated. Two out of three teacher variables interacted with the reference group effect of class-average intelligence. A complex pattern emerged for teachers' experience with giftedness: Some kinds of experience with giftedness lessened the influence of the reference group on students' probability of being nominated, whereas another kind strengthened the reference group effect, and again other experiences were not associated with it. When teachers held the belief that giftedness is domain-specific rather than holistic, the negative reference group effect was not as strong. Teachers' beliefs about the changeability of intelligence were not associated with students' probability of being nominated.

3.5 Discussion

We examined reference group effects of the class-average levels of intelligence on teachers' nominations of students for an enrichment program for gifted primary school students as well as how three giftedness-related teacher variables affected (the effects on) teachers' nominations. In line with previous research, our results indicated that, without knowing the intelligence scores of their students, teachers' decisions to nominate students for an enrichment program for gifted students were positively associated with students' fluid and crystallized intelligence. But whereas previous research has often been based on questionnaires (e.g., Endepohls-Ulpe & Ruf, 2005; Hany, 1997; Schack & Starko, 1990) or vignette approaches (e.g., Bianco et al., 2011), we were able to analyze the actual intelligence test scores of students who were nominated or not nominated for an enrichment program in a large sample, thus adding crucial information to the stream of research on nominating gifted students.

Most importantly, we found negative reference group effects: Students with the same intelligence scores, age, gender, and socioeconomic status from classes with higher average levels of intelligence had lower chances of being nominated than students from classes with lower average levels of intelligence. Our results indicate that teachers seem to use one of the most proximate reference groups (i.e., a student's class) to relativize their judgment of a student if they are provided with only a little information about their nomination task. Teachers' reliance on reference groups for their nominations is in line with many giftedness definitions (e.g., Subotnik et al., 2011) and checklist instructions (e.g., Pfeiffer & Jarosewich, 2003) that state that gifted individuals should be viewed in relation to a reference group (e.g., the average or other high-functioning students of the same age), and, furthermore, with recommendations to use local and subgroup norms for the identification of students for gifted education (e.g., Lohman, 2005; Lohman & Gambrell, 2012; Olszewski-Kubilius & Clarenbach, 2012).

However, the level on which teachers relativized their judgments for nominating students differed from the one set by the enrichment program. Our results indicate that teachers used a criterion that was relative on the class level, whereas the HCAP's intention was for teachers to apply a criterion that was relative on the population level (i.e., all students in Baden-Württemberg, Germany). Hence, on the one hand, teachers' use of the class as the reference group might provide valuable information for the identification process, for example, for programs that explicitly want to include students who are socioeconomically disadvantaged, have a minority status, or who speak another native language (see Lohman, 2005; Lohman & Gambrell, 2012). On the other hand, this orientation might be at odds with the criteria for identifying gifted students set by a specific program (e.g., the HCAP) and thereby needs to be addressed when asking teachers to nominate students.

Finally, our findings showed that two of the three teacher variables that we investigated were associated with the reference group effects and students' probability of being nominated. A complex pattern emerged such that different kinds of experience with giftedness had either positive, negative, or nonsignificant interactions with the reference group effect. These results show the importance of identifying the kind of experience teachers have with giftedness. For example, if someone gains experience only by learning about theoretical definitions of giftedness (e.g., by attending a lecture at a university), that person might view giftedness in comparison with a broad group of students (e.g., students of the same age). However, more in-depth research is needed to elaborate on what exactly causes different effects of the diverse kinds of experience with giftedness on reference group effects.

Further, the more teachers saw giftedness as domain-related, the less their nominations were affected by the negative reference group effect. One explanation might be that if giftedness is seen as holistic, fluid and crystallized intelligence are among the most important indicators and, therefore, the reference group's average intelligence level seems to matter. But if giftedness is seen as domain-related, even if intelligence is still an indicator of giftedness, the comparison with the class-average level of intelligence seems to be less important. In this case, teachers potentially emphasize domain-related abilities instead and, in turn, eventually compare those (narrower) abilities with a reference group.

In our analyses, we controlled for students' age, gender, and socioeconomic status. Our results that boys were more likely than girls to be nominated as gifted when students had the same fluid intelligence scores are in line with previous studies (Bianco et al., 2011; Endepohls-Ulpe & Ruf, 2005). However, a more complex pattern might be indicated because the effects were not found when students had the same crystallized intelligence scores (though the results were in the same direction). As found in several other studies (e.g., McBee, 2010; Speirs Neumeister et al., 2007), we also found that students with higher socioeconomic backgrounds had a higher probability of being nominated. Teachers nominated younger third-grade students more often than older ones. This result could be understood in the context of acceleration strategies for young gifted students and setbacks for students with (anticipated) difficulties. If children were identified as gifted, they may have entered school at an earlier age than their classmates or they may have skipped a grade. In a similar vein, if difficulties are anticipated or noticed, children might enter school one year later than their same-aged peers or have to repeat a school year.

Importantly, our models showed that a great deal of variance was still left unexplained. This result suggests that teachers use additional information to make their nomination decisions. In line with this, experts in the field of giftedness have often claimed that teachers should not be used as second guessers of intelligence, that their nominations include more than the judgment of students' intelligence, and that teachers can be helpful informants for programs that consider more than intelligence for their identifications (Foreman, & Gubbins, 2014; Hunsaker et al., 1997; Renzulli & Delcourt, 1986).

3.5.1 Strengths, limitations, and outlook

Our study has many strengths but also some limitations. As important strengths, we relied on a relatively large sample with data from 105 teachers and their 1,468 students from schools across an entire state. Although the schools in our study were chosen from a pool of schools that had previously nominated students for the HCAP, our sample should be rather representative overall in terms of the effects of reference groups on teachers' nominations for this program. Moreover, our design made it possible to combine students' individual characteristics, class-average-level characteristics, and teacher variables. Therefore, we did not have to rely on teachers' self-reports on crucial variables but were able to investigate reference group effects objectively.

We focused on third graders who were nominated for a program that provides enrichment to gifted primary school students. Whereas the interindividual stability of intelligence can be seen as quite high for this age group and beyond (e.g., Gottfried, Gottfried, & Wright, 2006; McCall, 1977; Rost, 2013; Schneider, Niklas, & Schmiedeler, 2014), further research might explore the generalizability of the effects we found to both other populations and other gifted education strategies such as acceleration.

Teachers received information about the HCAP (e.g., that the focus of the project was on STEM subjects) and the general goal of the HCAP to cater to the 10% most gifted students in the state of Baden-Württemberg, Germany. They were not provided with a rating scale or with specific advice on how to nominate students for the enrichment program. This procedure gave us a suitable environment to test whether (intelligence-related) reference group effects occurred with regard to teachers' nominations of gifted students when no specific instructions for the nomination were given. However, the information about the focus of the HCAP on STEM subjects might have influenced teachers' nomination decisions. Also teachers' knowledge about the goal of the HCAP to offer enrichment to the 10% most gifted students in Baden-Württemberg might have triggered teachers' general use of a reference group.

Our investigation focused on the impact of students' fluid and crystallized intelligence on teachers' nominations and included only a few other variables. Certainly, as previous research has already shown (e.g., Baudson & Preckel, 2013a; Endepohls-Ulpe & Ruf, 2005), other student characteristics such as motivation and social behavior seem to influence teachers' perceptions of whether students are gifted. Furthermore, the identification criteria set by enrichment programs must match the programs' goals, content, and target groups (e.g., Brody & Stanley, 2005). For general enrichment programs, general intelligence might be an important criterion, whereas enrichment programs for specific domains are more likely to be interested in specific student profiles with domain-specific abilities.

Future studies might explore the role of teacher variables in moderating the effect of the reference group we found on teachers' nominations. We started by looking at three teacher variables. However, especially experience with giftedness and the belief that giftedness is holistic or domain-specific might be investigated with more differentiated measurements. In a similar vein, student-related variables such as school grades, interest, and personality could also be considered as potential moderators. These variables might even result in reference group effects on teachers' nominations if they are investigated on the class or school level.

In our sample, nearly 20% of the students—an average of 2.74 students per class—were nominated for the enrichment program for gifted students, despite the program's goal of admitting only about the 10% most gifted students from each cohort. The most likely reason for this rather high number is that we selected schools—although we did so randomly—from a pool of schools that had nominated students in the previous school years. Although this necessarily increased students' overall likelihood of being nominated, the impact on our main analyses was most likely minor because we were primarily interested in students' *relative* likelihood of being nominated as a consequence of frame-of-reference effects.

3.5.2 Conclusion

Overall, our study extends the existing literature on both giftedness and reference group effects by showing negative reference group effects of average levels of fluid and crystallized intelligence computed at the class level on teachers' nominations for an enrichment program for gifted primary school students (in Germany). These effects were influenced in part by teacher variables. Consequently, next to individual student features, reference group effects also have to be considered when gifted students are identified by teachers both in theory and in practice. Program administrators should hence reflect on reference group effects in relation to the goals of their program as they follow the often-mentioned recommendations for selecting program participants (e.g., information and training of teachers, appropriate tools for nominations, and combinations of different methods and sources for selection). The program profile needs to be explicitly stated before the nomination process begins—especially with a focus on the reference groups that should be used.

4

Study 3:
A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students

Rothenbusch, S., Voss, T., Golle, J., & Zettler, I. (2016). A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students.
Manuscript submitted for publication.

Abstract

This study investigated teacher and parent ratings of teacher-nominated gifted elementary school students' verbal abilities, mathematical abilities, deductive reasoning, creative thinking, and engagement, and connected the ratings with school grades. We compared teacher and parent ratings concerning accuracy levels and halo effects. Further, we analyzed the congruence between teachers and parents and how teacher-parent congruence was related to school grades. Results were based on 294 teachers, 535 parents, and 408 elementary school students from an enrichment program. The results indicated that teachers and parents had the same accuracy levels, but teachers' ratings were more strongly affected by halo effects than parents' ratings. On average, the congruence between teachers and parents was low to moderate. Higher teacher and parent ratings were connected with better German grades in an additive manner. High parent ratings of mathematical abilities and engagement reduced the connection between teacher ratings and math grades.

Keywords: teacher ratings, parent ratings, accuracy, halo effects, teacher-parent congruence

A Comparison of Teacher and Parent Ratings of Teacher-Nominated Gifted Elementary School Students

4.1 Introduction

Teachers and parents are conjointly important for students' school performance (for reviews, see Christenson, 1999; Glueck & Reschly, 2014). Specifically, research has shown that the congruence between teacher and parent ratings of students as well as the congruence between teacher and parent expectations of students is linked to students' academic achievement (Brenner & Mistry, 2007; Peet, Powell, & O'Donnel, 1997). One of the most prominent questions in gifted education is how gifted students can transform their potential into excellent achievement (see Shavinina, 2009), illustrating the importance of the congruence between teacher and parent ratings for gifted students and gifted education, too. Although some studies have investigated teacher and parent ratings of students' cognitive abilities, creativity, and motivation (e.g., Geiser, Mandelman, Tan, & Grigorenko, 2016; Sommer, Fink, & Neubauer, 2008), all of which are seen as central facets of giftedness (see Sternberg & Davidson, 2005), studies have seldom concentrated on gifted students (e.g., Chan, 2000). Furthermore, empirical findings on the connection between teacher-parent congruence in ratings and gifted students' school achievement are sparse (Glueck & Reschly, 2014).

The contribution of the present study is thus twofold: First, this study contributes to the literature on the comparison of teacher and parent ratings of gifted students by presenting a comparison of teacher and parent ratings of German elementary school students who were nominated for a statewide enrichment program for gifted students by their teachers. At this juncture, we considered the following facets of giftedness: cognitive abilities (i.e., verbal abilities, mathematical abilities, and deductive reasoning), creative thinking, and engagement. Second, this study extends the literature by presenting an investigation of the effects of teacher-parent congruence in ratings on students' school grades in German and mathematics.

4.1.1 Gifted students

Modern scientific conceptions of giftedness (e.g., Heller, Perleth, & Lim, 2005; Subotnik, Olszewski-Kubilius, & Worrell, 2011, 2012) consider high general or domain-specific cognitive abilities to be the main aspects of giftedness but also list high creative, motivational, and beneficial environmental characteristics. Similarly, teachers' and parents' beliefs about giftedness typically comprise the idea that gifted students have high cognitive abilities, are creative, and are motivated to learn (Buckley, 1994; Endepohls-Ulpe & Ruf, 2005; Louis, 1992; Miller, 2009; Schack & Starko, 1990).

To identify gifted students, different methods such as tests, teacher nominations, or parent ratings as well as different criteria such as intelligence test scores or achievement outcomes are used in different combinations (e.g., Carman, 2013; National Association of Gifted Children, 2013). As a consequence, different groups of students are considered gifted. For example, a recent meta-analysis (Acar, Sen, & Cayirdag, 2016) distinguished nonperformance methods such as nominations and ratings by teachers and parents from performance methods that included tests of academic achievement, cognitive ability, and creativity. The nonperformance methods were moderately correlated with the performance methods (r = .30). Thus, teachers and parents saw a somewhat different group of students as gifted than one would have expected on the basis of ability and achievement tests.

4.1.2 Comparison of teacher and parent ratings of students' facets of giftedness

Teachers and parents have different perspectives on children (Harder, 2015; Petscher & Li, 2008). Typically, parents have known their children for longer and know their children better than teachers, whereas teachers—unless the parents are teachers themselves—have more professional knowledge in the area of education and academic development. Further, teachers, in comparison with parents, typically see a child in a larger number and variety of academically relevant situations, and they can compare a child with a larger group of other children.

Teachers' and parents' ratings of student characteristics that are seen as facets of giftedness (e.g., cognitive abilities, creativity, and motivation; see Sternberg & Davidson, 2005) have often been investigated separately (e.g., Baudson, Fischbach, & Preckel, 2014; Genser, Straser, & Garbe, 1981; Gralewski & Karwowski, 2013; Rennen-Allhoff, 1991; Skinner, Kindermann, & Furrer, 2009). A few studies have examined teacher and parent ratings together (e.g., Baudson & Preckel, 2013b; Geiser et al., 2016; Miller & David, 1987; Peet et al., 1997; Petscher & Li, 2007; Sommer et al., 2008) but did not concentrate—with a few exceptions (e.g., Chan, 2000)—on gifted students. In the following, we will summarize the main results across these studies with regard to the accuracy of teacher and parent ratings, halo effects in teacher and parent ratings, as well as teacher-parent congruence.

4.1.2.1 Comparison of the Accuracy of Teacher and Parent Ratings

Teachers and parents are often asked to rate different student characteristics (e.g., their mathematical abilities) on a Likert scale. The accuracy of these ratings is typically determined by comparing the ratings with the test scores or self-reports of the students (Machts, Kaiser, Schmidt, & Möller, 2016). Teachers and parents seem to be more accurate in rating students' cognitive abilities than in rating students' creativity (Machts et al., 2016; Schrader, 2010; Sommer et al., 2008; Urhahne, 2011). Teacher and parent ratings of cognitive ability were found to have similar average associations with test scores of general cognitive ability ($r_{\text{teachers}} = .56$, $r_{\text{parents}} = .50$; Sommer et al., 2008), but Geiser et al. (2016) reported a higher convergent validity concerning analytical abilities for teachers than parents. The correlations between ratings and test scores for students' verbal ($r_{\text{teachers}} = .57$, $r_{\text{parents}} = .19$), mathematical ($r_{\text{teachers}} = .70$, $r_{\text{parents}} = .43$), and figural abilities ($r_{\text{teachers}} = .53$, $r_{\text{parents}} = .35$) were higher for teachers than parents in a study by Miller and Davis (1992). However, on the basis of a comparison of teachers' and parents' difference scores (i.e., the difference between their ratings and students' test scores), the authors concluded that teachers' and parents' judgments were equally accurate. The association between ratings of creativity and students' test scores was higher for teachers than parents ($r_{\text{teachers}} = .34$, $r_{\text{parents}} = .24$) in Sommer et al.'s (2008) study, and Geiser et al. (2016) again reported a higher convergent validity for teacher ratings than parent ratings. Concerning students' motivation, associations between ratings and students' self-reports were reported to be low to medium for teachers ($r_{\text{teachers}} \leq .30$, as summarized by Spinath, 2005) and low or high for parents ($r_{\text{parents}} = .25$, Helmke & Schrader, 1989; $r_{\text{parents}} = .69$, Genser et al., 1981). Studies that have directly compared the accuracy of the two kinds of ratings of students' motivation have rarely been conducted.

4.1.2.2 Comparison of halo effects in teacher and parent ratings

The correlations between different characteristics that were rated by teachers or parents have often found to be higher than between corresponding student data such as tests or self-reports (e.g., Li, Lee, Pfeiffer, & Petscher, 2008; Sommer et al., 2008). For example, in a study by Urhahne (2011), correlations between teachers' ratings of students' mathematical abilities, creativity, and task commitment ranged from $r = .54$

(mathematical abilities and creativity) to r = .67 (mathematical abilities and task commitment), but correlations between students' test scores for mathematical abilities and creativity and self-reports for task commitment ranged from r = -.07 (mathematical abilities and task commitment) to r = .25 (mathematical abilities and creativity). This phenomenon has been termed the *halo effect* (e.g., Babad, Inbar, & Rosenthal, 1982; Burke, Haworth, & Ware, 1982; Mason, Gunersel, & Ney, 2014). Fisicaro and Lance (1990) offered three explanations for halo effects: first, a broad general impression of the person. Second, one highly salient factor might influences all ratings. Third, raters might be unable to conceptually discriminate between the different dimensions. On a descriptive level, the correlations between characteristics were higher for teacher ratings than for parent ratings (Chan, 2000; Petscher & Li, 2008; Pfeiffer et al., 2008), but the differences in the strengths of the correlations have, to the best of our knowledge, not yet been tested for significance.

4.1.2.3 Teacher-parent congruence in ratings of the same characteristic

When teacher and parent ratings of one characteristic have been directly compared with each other, studies have reported moderate to high correlations when the ratings were of students' cognitive abilities (Geiser et al., 2016; Miller & Davis, 1992; Sommer et al., 2008; Spinath & Spinath, 2005). On the other hand, however, weak associations were reported for teacher and parent ratings of students' creativity (Geiser et al., 2016; Runco, 1989; Sommer et al., 2008), whereas moderate correlations have been reported for students' motivational characteristics (Chan, 2000; Peet et al., 1997).

Overall, teachers might be similar or more accurate in rating students' cognitive abilities, more accurate in rating creativity, and similar or less accurate in rating motivation in comparison with parents. The literature indicates halo effects for both kinds of raters, but the effects might be stronger for teacher ratings. The congruence between teacher and parent ratings seems to fluctuate greatly in accordance with the characteristic in question. We inspected these facets of giftedness with regard to teacher-nominated gifted elementary school students. Furthermore, we were interested in the effect of congruence in teacher-parent ratings on students' school grades in German and mathematics.

4.1.3 Effects of congruence in teacher and parent ratings on students' school achievement

Scholars have stressed that students can perform better in school if the congruence between teachers and parents is high (e.g., for opportunities to learn, support, values, as well as standards and expectations that are realistic and high; for reviews, see Christenson, 1999; Glueck & Reschly, 2014). Glueck and Reschly (2014), however, noted that whereas the theoretical literature has often proposed a link between teacher-parent congruence and students' academic development, further empirical research is still needed. Indeed, only a few empirical studies have explored whether the congruence in teacher and parent ratings of student characteristics affects students' academic achievement. A cross-sectional study by Peet et al. (1997), for instance, showed that students' average school grades were better the more teachers and mothers agreed in their ratings of students' competence and engagement. The authors concentrated on the strength of congruence in teacher-parent ratings; however, stronger differentiations in the kind of high congruence (i.e., whether teachers and parents agree in a high or low rating of a characteristics) or low congruence (i.e., when teachers rate a characteristic higher than parents or the other way around) may be able to provide further insights. Support for the usefulness of such a detailed inspection comes from research on teacher and parent expectations regarding students' future academic achievement: In a cross-sectional study (Brenner & Mistry, 2007), teachers and parents were asked to indicate whether they expected students to finish high school or go to college. They found that teacher and parent expectations conjointly contributed to students' academic achievement in an additive manner. Students' academic achievement was highest when both raters had high expectations and lowest when both raters' expectations were low. However, the negative effect of low teacher expectations on student achievement was cushioned when mothers had high expectations. Corresponding research has yet to be conducted on gifted students. We herein aimed to tackle this issue.

4.2 Research Questions

We investigated teacher and parent ratings of elementary school students who were nominated as gifted by their teachers. We concentrated on ratings of five facets of giftedness: cognitive abilities (i.e., verbal and mathematical abilities, and deductive reasoning), creative thinking, and engagement. Specifically, on the basis of previous research, we pursued four research questions that were sorted into two objectives.

Objective 1 involved a comparison of teacher and parent ratings: (a) We compared teacher and parent ratings in relation to accuracy levels, (b) we did the same in relation to halo effects, and (c) we examined teacher-parent congruence in ratings of the same student characteristic. (a) We expected that the accuracy of ratings would be either similar or higher for teachers (Geiser et al., 2016; Miller & Davis, 1992; Sommer et al., 1992) concerning cognitive abilities, would be higher for teacher ratings in relation to creative thinking (Geiser et al., 2016; Sommer, 2008), and would not be different or would be lower for teachers concerning engagement (Genser et al., 1981; Helmke & Schrader, 1989; Spinath, 2005). (b) We hypothesized that we would find halo effects in the teacher and parent ratings (Li, Lee, Pfeiffer, & Petscher, 2008; Sommer et al., 2008) and that the correlations between teacher ratings of the five facets of giftedness would be higher than the correlations between the parent ratings (Chan, 2000; Petscher & Li, 2008; Pfeiffer et al., 2008). (c) We expected a moderate to high correlation between teacher and parent ratings of cognitive abilities (Miller & Davis, 1992; Sommer et al., 2008), a weak correlation between their ratings of creative thinking (Geiser et al., 2016), and a moderate correlation between their ratings of engagement (Chan, 2000).

Objective 2 involved effects of teacher-parent congruence in ratings on students' school grades: We investigated for each of the facets of giftedness whether the teacher-parent congruence in ratings was connected with students' school grades in German and mathematics. In relation to Brenner and Mistry (2007), we had two hypotheses: First, school grades should be best when teacher and parent ratings were congruently high and worst when teacher and parent ratings were congruently low. Second, we expected that when teacher ratings were low, high parent ratings were connected with better school grades than low parent ratings. This hypothesis was investigated through a moderation analysis. In support of the hypothesis, the analysis should show that high parent ratings reduced the association between teacher ratings and school grades.

Overall, our investigation thus provides information about teacher and parent ratings of facets of giftedness of teacher-nominated gifted students and extends previous research on the effects of teacher-parent congruence on students' school achievement.

Table 4.1

Sample Sizes in the Intersections Between the Three Data Sources

	Teacher ratings N	Parent ratings N	Tests and self-reports N
Teacher ratings	294		
Parent ratings	271	535	
Tests and self-reports	133	382	408

Note. $N_{\text{total}} = 572$; $N = 107$ for all three data sources.

4.3 Method

4.3.1 Participants

In 2010, an enrichment program for gifted elementary school students called the Hector Children's Academy Program (HCAP) was established in the German state of Baden-Württemberg. Sixty-one academies belong to the HCAP. These academies are located at elementary schools and offer enrichment classes for the 10% most gifted students in the state. About 60% of the classes cover STEM- (Science, Technology, Engineering, and Mathematics) related themes, whereas the remaining classes address other topics (e.g., art, languages). To participate in the HCAP, students have to be nominated by their teachers, and parents have to give their permission.

The data used in our study stem from a larger intervention study on the process of identifying gifted students. Herein, we used a subsample from this larger study for which we had information from at least two sources out of (a) teacher ratings, (b) parent ratings, (c) students' test data, or (d) students' self-report data. This resulted in data on 572 HCAP students from Grade 3 (43.53%) or Grade 4 (56.47%). The students (46.73% female; age: M = 9.25, SD = 0.65 years) came from 189 schools (M = 3.03, SD = 2.74, Min = 1, Max = 21). One hundred eighty-eight teachers rated 294 of the HCAP students on one to five rating scales (M = 1.58, SD = 0.93). For each of 535 HCAP students, we had one rating from the parents. Mothers and fathers rated the child together in 32.42% cases, only mothers in 60.81% cases, and only fathers in 5.49% cases (1.28% missing). The intersection between the three sources differed (see Table 4.1).

4.3.2 Procedure

The data for the present investigation came from a study that took place in the first half of the 2013-2014 school year. The study was voluntary for all participants (i.e., teachers, parents, and students). To ensure participants' anonymity, all data were collected and kept safe by the academy until the study was over. After the study, the data were coded, sorted, anonymized, and sent to the research team.

After information sessions about the intervention study, 12 of 60 academies agreed to participate. The study involved two phases. First, before the HCAP classes started, we collected rating scale data from the teachers who nominated the students for the HCAP. The nomination was a global and undifferentiated judgment of students and was executed by registering students with their parents' agreement at an academy. Teachers used the rating scales to rate the students whom they had nominated and handed the rating scales in with the registration. Teachers were informed that the goal of the HCAP is to cater to the 10% most gifted students, but otherwise, they were given no systematic information about giftedness. Teachers received no feedback from the academy about the nominated students. Students did not undergo further assessment. Most nominated students participate in the HCAP (e.g., r = .91 between nomination and participation; Rothenbusch, Zettler, Voss, Lösch, & Trautwein, 2016).

Second, during one of their classes, students were tested and questioned by trained test administrators. Students received an envelope from the test administrators with a questionnaire for their parents and were asked to pass it on to them. It was sent back to the academy in a sealed envelope (either through their children who passed it on to their HCAP instructors or by post in an enclosed stamped envelope).

4.3.3 Measures

We assessed students' verbal abilities, mathematical abilities, deductive reasoning, creative thinking, and engagement on the basis of teacher ratings (TRs), parent ratings (PRs), tests, and self-reports. Descriptive statistics can be found in Table 4.2.

4.3.3.1 TRs and PRs of student characteristics

TRs and PRs were examined with newly developed rating scales for the assessment of verbal and mathematical abilities, deductive reasoning, creative thinking, and engagement. Decisions about which content was measured were based on modern definitions of giftedness such as the ones by Renzulli (2005a) and Subotnik et al. (2011, 2012). The items for TRs and PRs had the same wording. However, the anchors of the 5-point scale varied for teachers and parents. Teachers rated whether students were *among the 5%, 10%, 25%, 50% best students of their age* (rated 5, 4, 3, or 2, respectively), or

Table 4.2

Descriptive Statistics for Teacher Ratings, Parent Ratings, and Student Data

	Teacher ratings[a]			Parent ratings[b]			Student data[c]		
	M	*SD*	*Min - Max*	*M*	*SD*	*Min - Max*	*M*	*SD*	*Min - Max*
Verbal abilities	3.77	1.02	1.00 – 5.00	4.40	0.75	1.67 – 5.00	54.86	16.27	8 – 92
Mathematical abilities	3.60	1.08	1.00 – 5.00	4.09	0.86	1.00 – 5.00	54.66	17.38	11 – 100
Deductive reasoning	3.65	0.99	1.00 – 5.00	4.13	0.67	1.50 – 5.00	76.13	22.93	12 – 100
Creative thinking	3.62	0.93	1.25 – 5.00	4.24	0.66	1.00 – 5.00	1.41	0.26	1.00 – 2.81
Engagement	3.93	0.98	1.25 – 5.00	4.12	0.69	2.00 – 5.00	3.43	0.48	1.31 – 4.00

[a] Estimations based on the unweighted items that loaded ≥ .50 on the factors. Value labels: *among the 5%* (rated 5), *10%* (rated 4), *25%* (rated 3), *50%* (rated 2) *best students of their age*, or *among the remaining 50%* (rated 1). Treated as interval-scaled. If treated as ordinal-scaled, Mdn of all five factors = 4.00, 25th percentile for all five factors = 3.00, 75th percentile for teacher ratings of verbal abilities and creative thinking = 4.50 and for teacher ratings of mathematical abilities, deductive thinking, and engagement = 5.00. [b] Estimations based on the unweighted items that loaded ≥ .50 on the factors. *Min* = 1 (*disagree*), *Max* = 5 (*agree*). [c] Tests for verbal abilities, mathematical abilities, and deductive reasoning: percent of correct answers, *Min* = 0%, *Max* = 100%. Test for creative thinking: *Min* = 1 (*not at all creative*), *Max* = 5 (*highly creative*). Self-report of engagement: *Min* = 1 (*strongly disagree*), *Max* = 4 (*strongly agree*).

among the remaining 50% (rated 1). As parents do not have access to the large reference group of students that teachers have for a comparison of students, parents' scale ranged from 1 (*disagree*) to 5 (*agree*). Further information about the factor structure is presented in the Results section. The wording of the items can be found in Table 4.3.

4.3.3.2 Student tests of cognitive abilities

Students' cognitive abilities were measured with three subtests from a German cognitive ability test for gifted third graders, namely, the *Kognitiver Fähigkeitstest für Hochbegabte* (Cognitive Ability Test for the Gifted, KFT-HB 3; Heller & Perleth, 2007). Due to time restrictions, we administered only one out of two existing subtests for each of the three abilities measured by the test: verbal, mathematical, and nonverbal. On the subtest for verbal abilities, students had to find words with similar meanings (α = .60 in our study). On the subtest for mathematical abilities, students had to form mathematical equations (α = .80 in our study). On the subtest for nonverbal abilities, students had to find the correct figure in relation to given ones (α = .91 in our study). We used this subtest as a proxy for deductive reasoning. Each student took two subtests and thus was measured on two of the three abilities. We had three combinations of subtests that were block randomized across HCAP classes: (a) verbal – mathematical (*N* = 138), (b) verbal – nonverbal (*N* = 145), and (c) nonverbal – mathematical (*N* = 125).

Table 4.3

Factor Loadings From the ESEM With Invariant Factor Loadings Between Teacher and Parent Ratings

	VA	MA	DR	CT	E	R^2	Items
						TRs/PRs	The student/My child… (TRs/PRs)
Verbal Abilities (VA)	**.97**	.02	.01	-.01	-.02	.94/.86	VA1: can express herself/himself verbally in different ways.
	.87	-.03	.07	.00	.09	.91/.73	VA2: formulates precisely.
	.88	.09	.01	.04	-.02	.91/.72	VA3: has a large vocabulary.
Mathematical Abilities (MA)	.00	**.84**	.14	.03	-.05	.90/.74	MA1: finds solutions for mathematical problems on her/his own.
	.04	**.95**	-.07	-.01	.08	.92/.76	MA2: solves mathematical problems reliably.
	-.02	**.82**	.14	.01	.03	.89/.76	MA3: masters complicated arithmetic operations.
Deductive Reasoning (DR)	.09	.16	**.50**	.04	.19	.77/.45	DR1: applies learned things independently in new contexts.
	.17	-.05	**.69**	.05	.08	.78/.51	DR2: compiles valid generalizations.
	-.02	.02	**.98**	-.04	.03	.94/.76	DR3: solves complex problems.
	.07	.05	**.79**	.08	-.05	.81/.61	DR4: understands abstract ideas.
Creative Thinking (CT)	.13	.00	-.01	**.71**	.10	.74/.46	CT1: has many good ideas on a specific theme.
	-.01	.08	.22	**.63**	.09	.79/.56	CT2: often finds extraordinarily good solutions to problems.
	.04	-.04	-.05	**.93**	.02	.84/.59	CT3: likes to come up with new things.
	-.02	.03	.08	**.96**	-.10	.87/.72	CT4: finds original solutions.
Engagement (E)	.03	.07	.16	.03	**.68**	.78/.45	E1: gets excited about new tasks.
	-.01	.07	.05	-.04	**.89**	.86/.56	E2: works persistently to solve a problem.
	.07	-.01	-.17	.02	**.93**	.78/.45	E3: works on tasks even when his/her initial endeavor is not successful.
	-.05	-.02	.05	.07	**.87**	.80/.51	E4: reacts enthusiastically when challenged.
Factor determinacy (TRs/PRs)	.99/.96	.98/.95	.98/.94	.97/.92	.97/.90		
Cronbach's alpha (TRs/PRs)[a]	.96/.90	.96/.89	.94/.83	.94/.83	.93/.78		

Note. TR = Teacher ratings; PR = Parent ratings. Results from M7.

[a] Cronbach's alpha based on the items that loaded ≥ .50 on the corresponding factor.

4.3.3.3 Student test of creative thinking

To estimate students' potential for creative thinking, we applied a verbal open-ended unusual uses task in which students had to produce creative examples for a common object. Within 2 min, students had to generate creative answers to the question "What can you do with a wooden board?" Students were instructed to "be creative" because research has indicated that such instructions increase the validity of the divergent thinking scores (O'Hara & Sternberg, 2001). As an indicator of creativity, we used the *average creativity* index (Silvia et al., 2008). Three raters rated the answers with respect to the criteria uncommonness, remoteness of associations, and cleverness on a 5-point Likert scale ranging from 1 (*not at all creative*) to 5 (*highly creative*). For every student and rater, the mean was calculated for the ratings of the students' answers. The resulting three scores for each student were averaged. The *ICC* between the three raters was .80.[1]

4.3.3.4 Student self-reports of engagement

We measured students' engagement with the *Anstrengungsbereitschaft* (willingness to make an effort) subscale from the German achievement motivation questionnaire *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen* (Questionnaire to Measure Emotional and Social School Experiences from Elementary School Third and Fourth Graders, FEESS 3-4, Rauer & Schuck, 2003). The subscale consists of 13 items ($\alpha = .79$ in our study) that are rated on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). An example item is "I also try to solve very difficult tasks."

4.3.3.5 Students' school grades

We ascertained students' report card grades for German ($M = 1.55$, $SD = 0.57$, $Min = 1$, $Max = 3$) and mathematics ($M = 1.38$, $SD = 0.54$, $Min = 1$, $Max = 3$) through the parent questionnaire. School grades can range from 1 (*excellent*) to 6 (*unsatisfactory*).

[1] We also calculated the quantitative creativity indices *fluency* (sum of all answers per student) and *uniqueness* (unique answers in the data set received a 1 and were then summed per student). In a factor analysis, the factor loadings of the three indices were low so that we could not justify a creativity factor. Therefore, we used the *average creativity* rating as the sole indicator of creativity in the study. The statistical significance of the associations of teacher and parent ratings of creativity with students' test scores were comparable across the three indices.

4.3.4 Analyses

We ran analyses using the Mplus 7.4 software package (Muthén & Muthén, 1998-2015). The indicators for the factors of TRs and PRs were treated as interval-scaled.[2] We used the MLR estimator and handled missing data with the full information maximum likelihood (FIML) algorithm, which uses the full information of the covariance matrices (Enders, 2010). Test and questionnaire scores were standardized separately for Grades 3 and 4, respectively, and then combined into one variable for the analyses. Because multiple significance testing increases the possibility of falsely rejecting the null hypothesis (i.e., inflated Type I error rate), we applied the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). Consequently, p-values $\leq .030$ were considered statistically significant at an overall level of $\alpha = .05$ in this study.

The levels of missing data within the three combinations of two data sources were as follows: (1) TRs and student data: $M = 7.25\%$, $SD = 11.82\%$, $Min = 0.75\%$ for five teacher-rated items, $Max = 36.84\%$ for the deductive reasoning test score; (2) PRs and student data: $M = 6.75\%$, $SD = 9.95\%$, $Min = 0.52\%$ for the engagement self-report, $Max = 34.55\%$ for the deductive reasoning test score; (3) TRs and PRs: $M = 2.77\%$, $SD = 1.36\%$, $Min = 0.74\%$ for a teacher-rated item, $Max = 7.75\%$ for students' math grade.

4.3.4.1 Hierarchical data structure

The teachers in our sample rated one or more students ($M = 1.58$, $SD = 0.93$). Because this dependence can lead to a violation of the assumption of traditional analysis approaches (e.g., OLS regression) that residuals are uncorrelated (Snijders & Bosker, 2012), we inspected *Intraclass Correlations* (ICC) that can range from 0 (total independence of observations from the cluster variable) to 1 (maximum dependence of observations from the cluster variable). In the present study, ICCs varied between .02 for one item from the PR scale for deductive reasoning and .58 for one item from the TR scale for engagement ($M = .25$, $SD = .20$). Even small ICCs of .05 or .01 can bias estimations in conventional OLS regression (Cohen, Cohen, West, & Aiken, 2003). Therefore, we applied the *type = complex* procedure in Mplus 7.4 to adjust the standard errors of the regression coefficients (for more information, see Muthén & Satorra, 1995).

---

[2] Estimations that were based on factors with interval-scaled or ordinal-scaled indicators were similar across the analyses.

4.3.4.2 Comparison of the factor structure of TRs and PRs

We conducted exploratory structural equation modeling (ESEM) because even small cross-loadings can inflate the correlations of CFA factors (Marsh et al., 2014). As recommended by Marsh et al. (2010; 2012), we used an oblique geomin rotation with an epsilon value of 0.5 (the default in Mplus 7.4). To examine the factor structure (see Table 4.4), we ran a sequence of five ESEM models, separately for TRs and PRs, with one to five factors (Models M1-5 for TRs and PRs). We assessed the fit of the models (see the passage on "goodness of fit" at the end of the Analyses section) and the factor loadings, following the rule of thumb (Cudeck, 2000; Tinsley & Tinsley, 1987) that items should have loadings of at least .30 on the target factor and should not have cross-loadings above .30 on any other factor. After establishing the factor structure of TRs and PRs, we calculated two ESEM models (M6 and M7) that included both the ESEM set for TRs and the ESEM set for PRs. In M6, the factor loadings were allowed to vary freely, and we investigated whether the items still had high loadings on the target factors and had low cross-loadings on other factors. To ensure the comparability of raters, we restricted the factor loadings between the raters to invariance (Meredith, 1993) in M7 and tested whether the model fit of M7 was not worse than the model fit of M6 (see the passage on "goodness of fit").

4.3.4.3 Objective 1: Comparison of TRs and PRs

Under the condition that the model fit of M7 was not worse than the model fit of M6, we added the test and self-report data as manifest variables to the ESEM sets with invariant factor loadings in M7 (see M8 in Table 4.5) to compare the associations between TRs, PRs, and students' tests and self-reports. To obtain correlations, the covariance estimates were standardized. To test whether the correlations differed significantly from each other, we used Fisher's z-transformations on the correlations and then calculated the difference between the correlations. We tested whether the differences were significantly different from 0.

4.3.4.4 Objective 2: Associations of teacher-parent congruence with students' school grades.

To test whether teacher-parent congruence was associated with students' school grades, we changed to the ESEM-within-CFA (EWC) framework because it overcomes some limitations of ESEM. In our case, we investigated the specific effects of two corresponding TR and PR factors per model and aimed to specify the latent interaction between the TR and PR of the same characteristic. This is not possible with ESEM factors in Mplus 7.4. More information on EWC can be found in an overview on ESEM by Marsh et al. (2014). For the EWC models, we used the parameter estimates from M7. Two EWC models per characteristic were estimated (thus, 10 models in total, see Table 4.6). The first model for each characteristic (Models VA1, MA1, DR1, CT1, and E1) included a regression analysis with school grades as the dependent variable and ratings, tests, and self-reports as predictors. To test for interaction effects, latent interaction terms for TRs and PRs were added to the regression in the second model for each characteristic (Models VA2, MA2, DR2, C2T, and E2).

4.3.4.5 Goodness of fit

We evaluated the model fit by computing the χ^2 goodness-of-fit statistic and the Root Mean Square Error of Approximation (RMSEA) for which values below .08, .05, and .01 indicate a mediocre, good, and excellent fit to the data (MacCallum, Browne, & Sugawara, 1996), respectively. Furthermore, we inspected the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), and the Standardized Root Mean Square Residual (SRMR). For the TLI and CFI, values above .90 indicate an acceptable fit and values above .95 an excellent fit to the data, and for the SRMR, values below .08 are considered to indicate a good model fit (Hu & Bentler, 1999).

To compare the relative fit between nested models in the analysis of the factor structure of TRs and PRs and the invariance of factor loadings between TRs and PRs, we used $\Delta\chi^2$ tests that should be nonsignificant if nested models are comparable in their fit to the data. We also used the guidelines proposed by Chen (2007): The CFI, RMSEA, and SRMR values should not change more than -.010, .015, and .030, respectively. Furthermore, no model fits were available in Mplus 7.4 for latent moderated structural equations (Muthén & Muthèn, 1998-2015). Therefore, for Objective 2, after inspecting the model fit indices in the models without latent interaction terms, log-likelihood ratio tests were used to determine whether the more parsimonious models (models without a latent interaction term) showed significant losses in fit in relation to the more complex models (models with the latent interaction term). A significant log-likelihood ratio test would indicate that the model without the interaction term showed a significant loss in fit

Table 4.4

Factor Structure of Teacher and Parent Ratings

Models	χ^2	*df*	CFI	TLI	RMSEA	SRMR
			Teacher ratings (*N* = 294)			
M1 for TRs: 1 factor	1830.384^{***}	135	.663	.618	.207	.089
M2 for TRs: 2 factors	1208.935^{***}	118	.783	.719	.177	.073
M3 for TRs: 3 factors	701.658^{***}	102	.881	.821	.141	.047
M4 for TRs: 4 factors	273.937^{***}	87	.963	.935	.085	.018
M5 for TRs: 5 factors	90.208	73	.997	.993	.028	.007
			Parent ratings (*N* = 535)			
M1 for PRs: 1 factor	2207.594^{***}	135	.511	.445	.169	.126
M2 for PRs: 2 factors	1384.821^{***}	118	.701	.612	.142	.096
M3 for PRs: 3 factors	698.752^{***}	102	.859	.789	.105	.062
M4 for PRs: 4 factors	380.001^{***}	87	.931	.878	.079	.032
M5 for PRs: 5 factors	158.438^{***}	73	.980	.958	.047	.018
			MTMM models of teacher and parent ratings			
M6: 2x5 factors (FL unrestricted for raters)	649.389^{***}	445	.981	.973	.029	.026
M7: 2x5 factors (FL restricted for raters)	744.223^{***}	510	.978	.973	.029	.031

Note. TR = Teacher ratings; PR = Parent ratings. CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; FL = Factor loadings. Δ = Difference calculations in relation to M6.

in comparison with the model with the interaction term. A nonsignificant result would show that the model without the interaction term had no significant loss in fit in comparison with the model with the interaction term. Conclusions about whether the fit of the model with the interaction term was equal to or worse than the more parsimonious model could not be drawn (Maslowsky, Jager, & Hemken, 2015).

4.4 Results

4.4.1 Data preparation: Structure of teacher and parent ratings

Separately for the teacher ratings (TRs) and parent ratings (PRs), we tested the factor structure of the ratings with a sequence of ESEM models that assumed one to five factors. We expected to find a five-factor structure for TRs and PRs (i.e., one factor for each rated characteristic). The model fit indices are given in Table 4.4. Models with fewer than five factors (M1-M4 for TRs and PRs) did not meet the fit criteria and had worse model fits than the five-factor models (M5 for TRs and PRs; all *p*s < .05 for $\Delta\chi^2$ of M1-

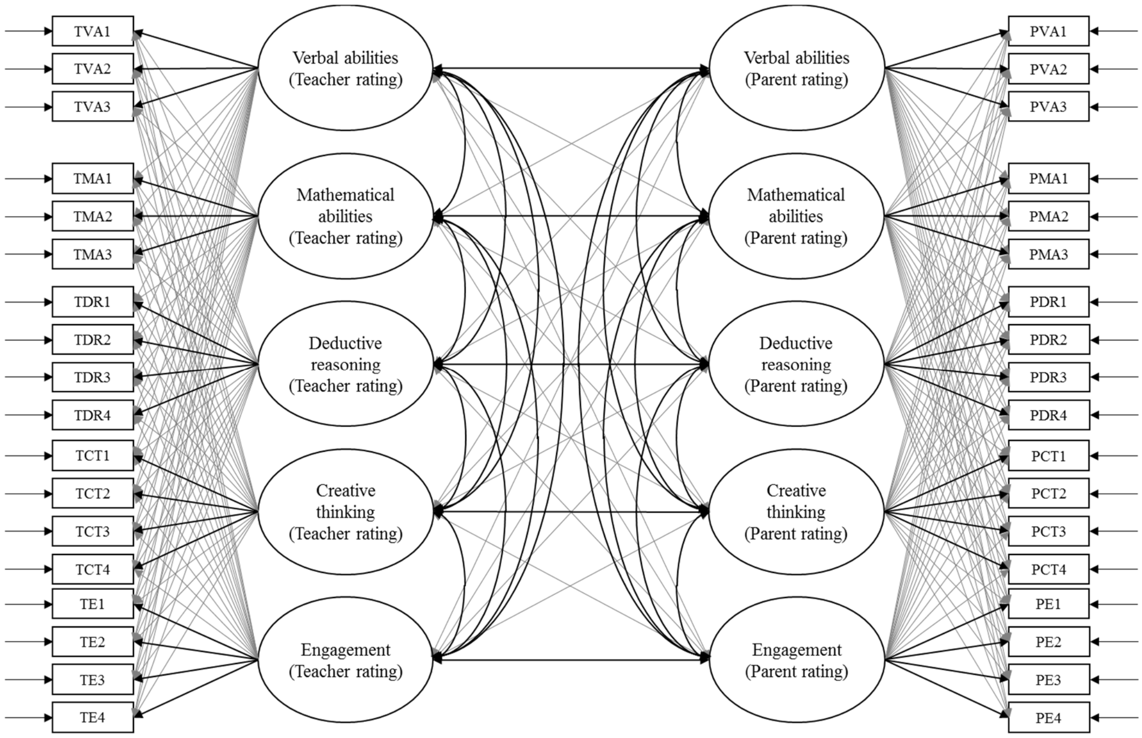

Figure 4.1. Exploratory structural equation model (ESEM) of teacher and parent ratings. Latent factors in circles, items in boxes. Main item loadings on factors, monotrait-heteromethod and heterotrait-monomethod correlations are highlighted in black and bold.

M4 vs. M5). The five-factor models had acceptable fits to the data. All items had loadings of at least .55 (mean loading = .84) for TRs and .41 (mean loading = .73) for PRs on the intended factors, and the cross-loadings did not exceed .17 for TRs or .27 for PRs. The correlations between the dimensions were high for TRs (mean r = .64, SD = .09) and medium for PRs (mean r = .34, SD = .15).

We combined the ESEM sets for TRs and PRs into one model (see Figure 4.1 and Table 4.4). In M6, all parameters were allowed to vary freely to test for factor-form invariance. The inspection of the model's fit indices showed a good fit to the data. In M7, we set the factor loadings to invariance between the raters. The model fit did not get worse, $\Delta\chi^2(65) = 77.226$, $p > .05$, indicating comparable factor structures for TRs and PRs. Details on the items and factor loadings of M7 are presented in Table 4.3.

4.4.2 Objective 1: Comparison of TRs and PRs

For Objective 1, we compared TRs and PRs in relation to their accuracy levels, explored whether they were affected by halo effects, and investigated the congruence of TRs and PRs. Therefore, we entered students' test scores and self-report data into the

Table 4.5

Associations Between Teacher Ratings, Parent Ratings, and Student Data

	Verbal abilities		Mathematical abilities		Deductive reasoning		Creative thinking		Engagement	
	r	*SE*	*r*	*SE*	*r*	*SE*	*r*	*SE*	*r*	*SE*
	Teacher ratings									
1. Teacher rating										
Mathematical abilities	.49***	.06								
Deductive reasoning	.68***	.04	.79***	.03						
Creative thinking	.68***	.04	.58***	.05	.74***	.04				
Engagement	.58***	.05	.60***	.06	.66***	.05	.65***	.05		
2. Parent rating										
Verbal abilities	.31***	.06	.11	.07	.07	.07	.07	.08	.22**	.07
Mathematical abilities	.04	.06	.40***	.06	.12	.06	.01	.07	.16*	.06
Deductive reasoning	.17**	.06	.30***	.06	.12	.07	.16*	.07	.25***	.06
Creative thinking	.10	.06	.20**	.06	.01	.07	.10	.07	.17*	.07
Engagement	.14*	.06	.11	.08	.04	.07	.11	.07	.28***	.06
3. Student data										
Verbal abilities	.35***	.09	.33***	.08	.45***	.08	.25*	.10	.22*	.09
Mathematical abilities	.15	.08	.26**	.09	.18	.10	.08	.11	.04	.11
Deductive reasoning	.22	.12	.27**	.10	.24*	.11	.17	.11	-.03	.10
Creative thinking	.02	.10	-.05	.08	.05	.10	-.05	.09	.08	.10
Engagement	.12	.10	.00	.09	.05	.09	.01	.10	.15	.10
	Parent ratings									
4. Parent rating										
Mathematical abilities	.10	.06								
Deductive reasoning	.35***	.05	.58***	.05						
Creative thinking	.24***	.05	.17**	.06	.48***	.06				
Engagement	.28***	.06	.42***	.06	.44***	.05	.29***	.07		
5. Student data										
Verbal abilities	.31***	.05	.22**	.07	.20**	.07	.08	.06	.11	.07
Mathematical abilities	-.01	.07	.22**	.07	.08	.07	-.08	.07	.13	.07
Deductive reasoning	.18**	.06	.34***	.07	.18**	.06	.04	.07	.25***	.06
Creative thinking	.05	.06	-.08	.06	.04	.05	.04	.05	.04	.06
Engagement	.09	.05	.18**	.06	.10	.06	.09	.06	.37***	.06
	Student data									
6. Student data										
Mathematical abilities	.27***	.07								
Deductive reasoning	.14	.09	.20*	.09						
Creative thinking	.02	.06	-.03	.06	-.02	.06				
Engagement	.10	.06	.12	.06	.06	.05	.08	.04		

Note: Results from M8. Based on the adjustment of significance tests with the Benjamini-Hochberg procedure (1995), *p*-values ≤ .030 are considered statistically significant at an overall level of $\alpha = .05$.
$^{*}p \leq .030$. $^{**}p < .010$. $^{***}p < .001$.

rating factors of M7 (i.e., M8). M8 had a good fit to the data, $\chi^2(640) = 938.462$, RMSEA = .029, CFI = .973, TLI = .965, SRMR = .034. All correlations between teacher ratings, parent ratings, and student data can be found in Table 4.5.

4.4.2.1 Differences in the accuracy of TRs and PRs

We investigated differences in the accuracy of TRs and PRs in relation to the corresponding student data. Table 4.5 presents the correlations between the ratings and the student data in Sections 3 and 5. TRs ($r_{TR} = .35, p < .01$) and PRs ($r_{PR} = .31, p < .01$) were moderately associated with the student data for verbal ability, weakly associated for mathematical ability ($r_{TR} = .26, p < .01$; $r_{PR} = .22, p < .01$) and deductive reasoning ($r_{TR} = .24, p = .01$; $r_{PR} = .18, p < .01$), and not significantly associated for creative thinking ($r_{TR} = -.05$, *ns*; $r_{PR} = .04$, *ns*). For students' engagement, the TRs were not significantly associated ($r_{TR} = .15$, *ns*) and the PRs were moderately correlated ($r_{PR} = .37, p < .01$) with students' self-reports. The TRs and PRs did not differ significantly from each other in their associations with the students' data (all *p*s > .030). The results of nondifferences between TRs and PRs were as expected for verbal abilities, mathematical abilities, deductive reasoning, and engagement but not for creative thinking for which we had expected higher TRs.

4.4.2.2 Difference in halo effects in TRs and PRs

We hypothesized that the ratings of different characteristics would be more highly intercorrelated within TRs and within PRs than within the corresponding student data (i.e., halo effects). Furthermore, we expected that TRs of different characteristics would be more strongly intercorrelated than PRs of different characteristics. The correlations within the three methods (i.e., teachers, parents, and students) can be found in Table 4.5 in Sections 1, 4, and 6. A comparison of the correlations within the TRs, within the PRs, and within the student data revealed that all correlations within the TRs were significantly higher than within the student data (see *TR – S* in Figure 4.2). In nine out of 10 cases, the correlations within the PRs were significantly higher than within the student data (see *PR – S* in Figure 4.2). Furthermore, correlations within the TRs were significantly higher than within the PRs in all cases (see *TR – PR* in Figure 4.2). Therefore, the results indicate that the ratings—TRs more strongly than PRs—were affected by halo effects.

4.4.2.3 Congruence between TRs and PRs

We examined the TR-PR congruence for the same characteristics. The correlations are depicted in Table 4.5 on the diagonal of Section 2. TRs and PRs were moderately correlated for verbal abilities and mathematical abilities ($r_{\mathrm{VA}} = .31, p < .01$; $r_{\mathrm{MA}} = .40, p < .01$), not significantly correlated for deductive reasoning and creative thinking ($r_{\mathrm{DR}} = .12$, *ns*; $r_{\mathrm{CT}} = .10$, *ns*), and weakly correlated for engagement ($r_{\mathrm{E}} = .28, p < .01$). The correlations between ratings of verbal abilities, mathematical abilities, and creative

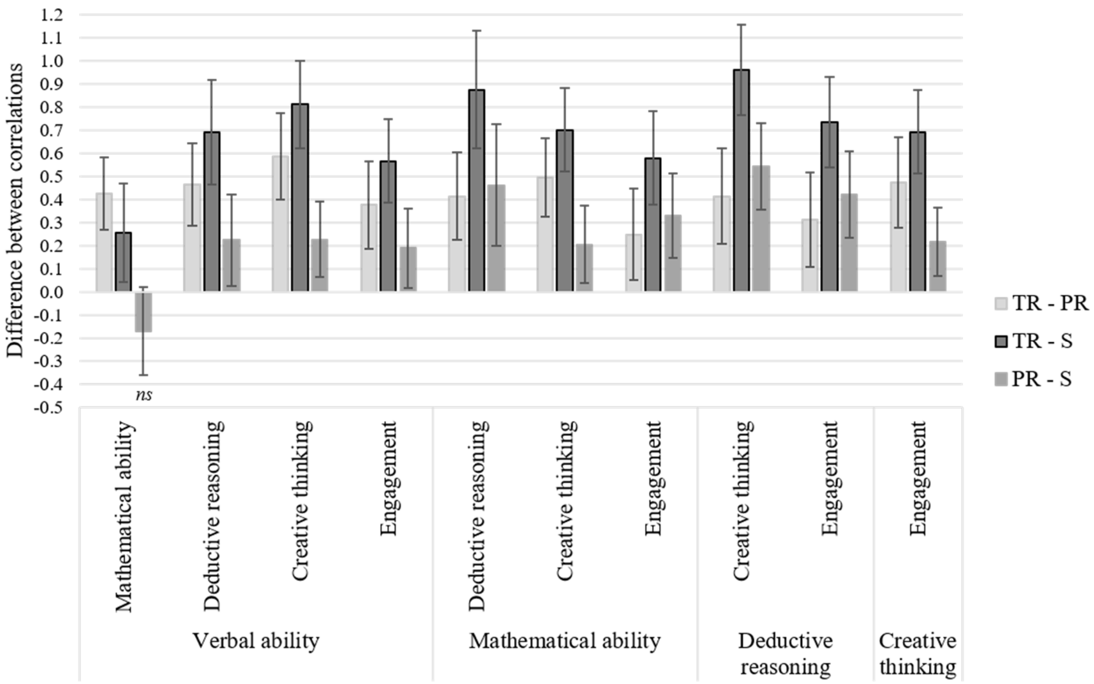

Figure 4.2. Differences between the corresponding heterotrait-monomethod correlations of teacher ratings and parent ratings (TR – PR), teacher ratings and student data (TR – S), and parent ratings and student data (PR – S) are depicted. Difference calculations are based on the *Fisher-z*-transformed correlations from M8. Error bars are 95% CIs. *ns* = the difference between the correlations was not statistically significant on the basis of the adjustment of significance tests with the Benjamini-Hochberg procedure (1995) at an overall level of $\alpha = .05$.

thinking were as expected, but the correlation between the ratings of deductive reasoning and engagement was lower than anticipated.

4.4.3 Objective 2: Effects of teacher-parent congruence in ratings on school grades

We examined whether TRs and PRs were conjointly associated with students' German and math grades. Thereby, lower values for school grades indicate better school grades. We ran two models for each characteristic (see Table 4.6)—one model to inspect whether TRs and PRs were significant predictors of students' school grades and one model to test whether the latent interactions between TRs and PRs affected school grades.

Higher TRs and PRs of verbal ability (Models VA1-2) were additively associated with better German grades (lower values). Therefore, for students with the same test scores for verbal abilities, German grades were best (lowest values) if teachers and parents both had high ratings and worst if both ratings were low. The interaction was not significant.

Table 4.6

Associations of Teacher and Parent Ratings With Students' School Grades While Controlling for Students' Test Scores and Self-Reports

	German Grade		Math Grade		German Grade		Math Grade	
	β	*SE*	β	*SE*	β	*SE*	β	*SE*
	Model VA1				Model VA2			
Verbal abilities *Teachers*	-.32 ***	.08			-.32 ***	.08		
Verbal abilities *Parents*	-.18 ***	.05			-.15 **	.05		
Interaction *Teachers* x *Parents*					.10	.05		
Verbal abilities *Students*	-.09	.11			-.09	.10		
R^2	.21 ***	.05			.20 ***	.05		
	Model MA1				Model MA2			
Mathematical abilities *Teachers*			-.22 **	.07			-.25 ***	.07
Mathematical abilities *Parents*			-.42 ***	.05			-.37 ***	.06
Interaction *Teachers* x *Parents*							.14 **	.05
Mathematical abilities *Students*			-.14	.07			-.16 *	.07
R^2			.36 ***	.05			.37 ***	.05
	Model DR1				Model DR2			
Deductive reasoning *Teachers*	-.27 ***	.07	-.49 ***	.08	-.27 ***	.07	-.48 ***	.08
Deductive reasoning *Parents*	-.14 *	.06	-.12	.06	-.15 **	.06	-.14	.07
Interaction *Teachers* x *Parents*					.04	.06	.11	.06
Deductive reasoning *Students*	-.16	.07	-.37 ***	.08	-.14	.08	-.33 **	.11
R^2	.12 *	.05	.37 ***	.07	.12 *	.05	.37 ***	.07
	Model CT1				Model CT2			
Creative thinking *Teachers*	-.26 **	.09	-.45 **	.14	-.27 **	.09	-.47 ***	.13
Creative thinking *Parents*	-.14 *	.06	.00	.06	-.14 *	.06	-.01	.06
Interaction *Teachers* x *Parents*					-.09	.08	.00	.06
Creative thinking *Students*	.07	.14	.21	.21	.09	.14	.23	.19
R^2	.09	.05	.22	.13	.10	.05	.23	.13
	Model E1				Model E2			
Engagement *Teachers*	-.23 **	.07	-.16	.15	-.22 **	.07	-.19	.12
Engagement *Parents*	-.37 ***	.06	-.39 ***	.08	-.34 ***	.06	-.35 ***	.08
Interaction *Teachers* x *Parents*					.10	.05	.16 **	.05
Engagement *Students*	-.01	.08	-.01	.08	-.03	.07	-.01	.08
R^2	.23 ***	.05	.21 ***	.06	.22 ***	.05	.22 ***	.06

Note: Lower values for school grades indicate better school grades. Based on the adjustment of significance tests with the Benjamini-Hochberg procedure (1995), p-values ≤ .030 are considered statistically significant at an overall level of α = .05. Model fit indices for the models VA1, MA1, DR1, CT1, and E1: RMSEA = .028-.037, CFI = .957-.977, TLI = .947-.972, SRMR = .032-.059. Based on log-likelihood ratio tests between the first and second model per characteristic, E2 fit better than E1; D (2) = 27.063, $p < .001$; and MA2 fit better than MA1; D (1) = 5.010, $p = .025$. VA1, DR1, and CT1 did not result in a significant loss in fit relative to VA2, DR2, and CT2; $D_{\text{VA1 vs.VA2}}$ (1) = 1.280, $p > .05$; $D_{\text{DR1 vs.DR2}}$ (2) = 3.684, $p > .05$; $D_{\text{CT1 vs.Ct2}}$ (2) = 2.321, $p > .05$.
$^{*}p \leq .030$. $^{**}p < .010$. $^{***}p < .001$.

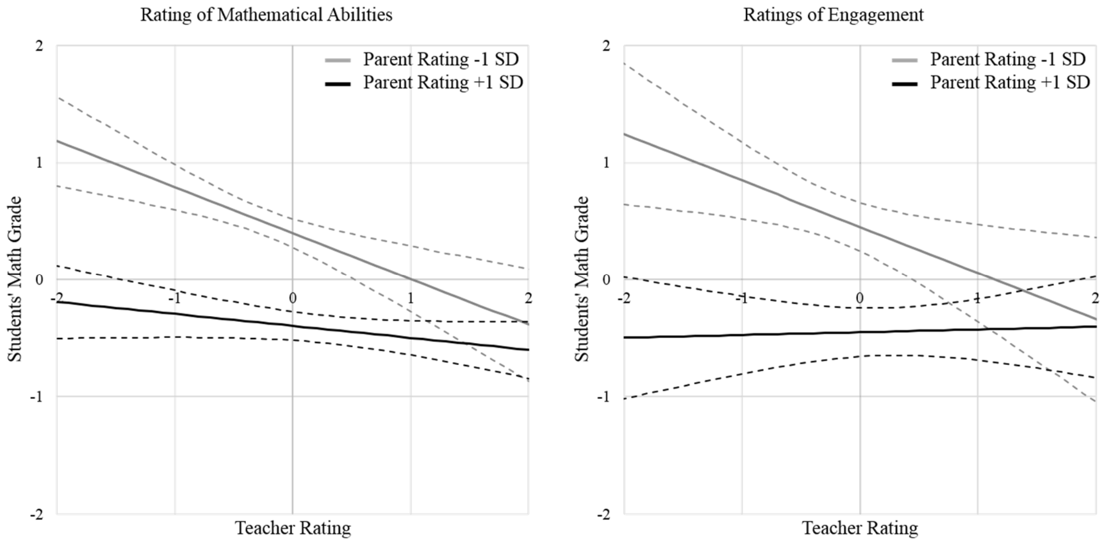

Figure 4.3. Effects of interactions between teacher ratings and parent ratings on students' math grades. Lower values for math grades indicate better math grades. The dotted lines are 95% CIs.

For the ratings of mathematical abilities while controlling for students' test scores for mathematical abilities (Models MA1-2, see also Figure 4.3), high congruence in high teacher-parent ratings were connected to better grades (lower values), and high congruence in low teacher-parent ratings were connected to worse grades (higher values). However, the significant interaction term indicated a multiplicative effect of TRs and PRs on math grades. When parents had low ratings, the association between TRs and math grades was diminished resulting in better math grades when TRs were low but PRs were high instead of congruently low TRs and PRs.

For deductive reasoning (Models DR1-2) and creative thinking (Models CT1-2), TRs and PRs were significantly associated with students' grades in German in an additive manner, but only TRs were significantly connected to students' math grade. The interaction was not significant. TRs and PRs of engagement (Models E1-2) were significantly and additively connected to students' grades in German. For students' math grades, the interaction between TRs and PRs of engagement was significant, indicating a pattern that was similar to the one for teacher-parent congruence in ratings of mathematical abilities (see Figure 4.3).

Overall, with the exception of the effects of the ratings of creative thinking and deductive reasoning on math grades, high congruence in high teacher-parent ratings were associated with better grades and high congruence in low teacher-parent ratings with worse grades. For ratings of mathematical ability and engagement, high PRs buffered the effect of TRs on students' math grades.

4.5 Discussion

With this study, we looked at teacher and parent ratings of facets of giftedness of elementary school students who received teacher nominations for a statewide enrichment program. Our study pursued two objectives: First, we compared teacher and parent ratings of students' cognitive abilities (i.e., verbal and mathematical abilities and deductive reasoning), creative thinking, and engagement with each other with regard to accuracy levels and halo effects. Further, we examined the congruence in teacher and parent ratings of the same characteristic. Second, we analyzed how the congruence in teacher-parent ratings of the same characteristic was associated with students' German and math grades.

4.5.1 Comparison of teacher and parent ratings of teacher-nominated elementary school students

4.5.1.1 Difference in the accuracy of teacher and parent ratings

Most previous work on teacher and parent ratings of cognitive abilities, creativity, and motivation was based on students from general education settings (e.g., Gagné & St Père, 2001; Geiser et al., 2016; Sommer et al., 2008), but seldom based on gifted students (cf. Chan, 2000, for gifted secondary school students). In our study with ratings of teacher-nominated gifted elementary school teachers, the accuracy of teacher ratings of all five facets of giftedness did not differ significantly from that of parents. Therefore, although teachers and parents had different perspectives on students (Harder, 2015; Petscher & Li, 2008), they seemed to be equally able to rate facets of giftedness.

Concerning cognitive abilities, the lack of difference was in line with Sommer et al. (2008) although Miller and Davis (1992) reported higher correlations between teacher ratings and test scores for verbal, mathematical, and figural abilities. For ratings of creativity, higher accuracy levels for teacher ratings were reported (Geiser et al., 2016; Sommer et al., 2008). A reason why teachers were not more accurate than parents in rating teacher-nominated gifted students might be that parents' but not teachers' accuracy in rating students' cognitive abilities was found to be higher if the children had high achievement levels and school grades (Miller & Davis, 1992). The authors offered several explanations such as that parents who can rate their children more accurately might be promoting their children more adequately resulting in more competent children (i.e., match hypothesis; Hunt & Paraskevopoulos, 1980) or that higher achieving children might be easier to detect for parents. Furthermore, parents' knowledge of their children's status as gifted (as determined by the teachers) might have resulted in the parents paying greater attention to their children's levels on these characteristics as the characteristics were possibly connected with the children's status as gifted. However, in comparison to other studies (e.g., Machts et al., 2016; Miller & Davis, 1992; Sommer et al., 2008; Urhahne, 2011), teachers and parents' accuracy level of rating cognitive abilities and creative thinking was mainly lower in this study. The teacher nominations might have led to a sample with a limited variance in these characteristics in comparison to samples from general classrooms, therefore the correlations might be artificially reduced due to statistical reasons (Wild, 1993). However, teachers' decision to nominate a student as gifted and parents' perceptions of their children's nomination status might have differently induced beliefs about how these students should be (e.g., Baudson, 2016; Baudson & Preckel, 2016) and thereby distorted teachers and parents' ratings. Studies are needed that include teacher and parent ratings of students with and without a gifted status to explore the reasons for the heights of accuracy levels and the (non)differences between teacher and parent ratings.

In line with previous studies, parent ratings of engagement were moderately connected to students' self-reports (Helmke & Schrader, 1989), whereas teacher ratings correlated weakly and not significantly with them (Spinath, 2005). Although the difference between teacher and parent ratings was not significant in our study, this result should be investigated further. It might indicate that parents can rate aspects of students' engagement better than teachers can, maybe due to parents' chances of observing their children out of school in learning-relevant situations such as when they do homework, prepare for tests, or talk about school and learning content.

4.5.1.2. Difference in halo effects in teacher and parent ratings

In agreement with previous research that indicated halo effects (e.g., Li et al., 2008; Peters & Gentry, 2010; Petscher & Li, 2008; Urhahne, 2011), cognitive abilities, creative thinking, and engagement were more strongly correlated when rated by teachers than when rated by parents or when based on student data (i.e., test scores and self-reports). In nine out of 10 cases, parent ratings were more strongly intercorrelated than the student data were. Therefore, teachers and parents connected the facets of giftedness more strongly than the student data implied. The ratings indicated rather homogeneous profiles of teacher-nominated students, whereas student data showed a high diversity of combinations of strengths and weaknesses on the characteristics. Our results exclude one of the reasons for halo effects mentioned by Fisicaro and Lance (1990): Our factor analyses indicated that both groups of raters were able to discriminate between the rated characteristics. The results might be related to teachers' and parents' belief that gifted students are, for example, highly intelligent, creative, and committed to learning (e.g., Buckley, 1994; Endepohls-Ulpe & Ruf, 2005; Louis, 1992). On the basis of this belief, they might have inferred rather similar levels of these characteristics. However, halo effects have also been observed for ratings of students who were not labeled gifted (e.g., Urhahne, 2011). In teachers, Anders, McElvany, and Baumert (2010) found a global factor for academic achievement that included cognitive abilities and motivation. Such a global factor might affect (to different extents) teacher and parent ratings of learning and achievement-relevant characteristics, regardless of students' level. The factor might be strongly connected to beliefs about giftedness at the highest level. Next to the necessity to investigate these possibilities, research should explore whether teachers' responsibility to nominate students might be connected to the higher correlations within the teacher ratings than within the parent ratings. Also in-depth analyses are needed to examine whether different profiles are hidden behind the high but not perfect correlations.

4.5.1.3 Congruence between teacher and parent ratings

The teachers and parents agreed that the students in this study should participate in a statewide enrichment program. Although this agreement might have been accompanied by a higher congruence in the ratings of facets of giftedness, congruence was not higher than the reported congruence levels in general classrooms. In alignment with previous research, teacher and parent ratings were moderately associated for verbal and mathematical abilities (Miller & Davis, 1992) and weakly correlated for creative thinking (Geiser et al., 2016). Teacher-parent congruence for ratings of deductive reasoning and engagement was lower than expected (Chan, 2000; Miller & Davis, 1992). The low to at most mediocre associations between teacher and parent ratings might be connected to relatively lower observability of the item content (e.g., "understands abstract ideas"), high subjectivity (e.g., "good ideas" and "original solutions"), or the differences in the observable situations for teachers and parents.

4.5.2 The connection of teacher-parent congruence with school grades

Our study offers important empirical support for the hypothesis that teacher-parent congruence matters for students' academic development (Glueck & Reschly, 2014). Specifically, we extended Peet et al.'s (1997) findings by showing that high congruence between teacher and parent ratings of facets of giftedness is not always associated with high student achievement in school. As expected from the research showing that congruence between teachers' and parents' expectations about students' achievement is connected with students' future school success (Brenner & Mistry, 2007), congruently high teacher-parent ratings of a facet of giftedness were mostly associated with better school grades, and congruently low teacher-parent ratings of a facet of giftedness were mostly connected with worse school grades—even when students had the same test or self-reported scores. This result has special importance as it emphasizes that teacher and parent ratings should be considered in conjunction, for instance, in attempts to transform high ability into high achievement. Therefore, the reasons for the at most mediocre congruence between teacher and parent ratings as well as ways to enhance teacher-parent congruence in rating students' strengths should be given more attention.

Furthermore, in accordance with the results on expectations presented by Brenner and Mistry (2007), the effects of teacher ratings of mathematical abilities and engagement on students' math grades were diminished if the corresponding parent ratings were high. Thereby math grades were not as worse when teacher ratings were low but parent ratings were high instead of congruently low teacher and parent ratings. Therefore, as an extension of Helmke's (2012) notion that a slight overestimation of students' abilities by teachers is educationally beneficial, our results indicate that parents' ratings of their children's abilities and engagement can also be an asset if the parents' ratings are higher than teachers' ratings—even for students who were nominated as gifted by teachers and who probably thus had rather good school grades already.

However, the variance explained in school grades was rather low, showing that teacher-parent congruence in ratings is a piece, but a rather small piece, of the puzzle of what explains students' school grades. As mentioned, the pattern we found is similar to the pattern reported by Brenner and Mistry (2007) for expectations. Ratings of student characteristics inform expectations of students' future characteristics, but expectations can also guide future ratings (Rubie-Davis, 2010; Timmermans, de Boer, & van der Werf, 2016). Hence, in-depth analyses are needed to trace the path between teacher-parent congruence-dissonance in ratings and in expectations. The next steps should be to explore the black box of how teacher-parent congruence-dissonance affects students' school grades and the other way around. In doing so, the role of actual versus perceived congruence might be examined.

4.5.3 Strengths and limitations

The presented results should be seen in the light of certain strengths and limitations of our study. First, we combined three important perspectives on facets of giftedness: teacher ratings, parent ratings, and student data (i.e., tests and self-reports). In addition, we measured a comprehensive range of characteristics. Therefore, we were able to investigate how ratings of different characteristics were connected and to set previous results on teacher and parent ratings of facets of giftedness in relation to each other.

Second, with our focus on teacher-nominated elementary school students, we contribute to the research on teacher and parent ratings, which has mostly been conducted on students from general classrooms (e.g., Geiser et al., 2016; Sommer et al., 2008). However, our group might differ in part from students who were identified as gifted on the basis of ability and achievement tests (Acar et al., 2016). Future research is thus needed to compare ratings of students who are seen as gifted by their teachers with students who are labeled as gifted on the basis of their test scores. Furthermore, we focused on elementary school students, and our study was conducted in a German state. Students came from many different elementary schools in that state, so the results might be rather representative of teacher and parent ratings of teacher-nominated elementary school students in this area, but the generalizability to ratings of, for example, teacher-nominated students of another age or nationality must be scrutinized.

Third, our results were based on the ESEM and EWC approaches, which yield more accurate estimations of correlations than CFA (Marsh et al., 2014), although their interpretation might be more difficult because they allow cross-loadings. Fourth, teachers simultaneously rated and nominated students for the HCAP, whereas parents rated children who participated in the HCAP. Therefore, teachers and parents did not have the same conditions for their ratings. Future research should investigate the extent to which these different conditions affect ratings.

Fifth, we used report card grades for German and mathematics as indicators of students' academic achievement. School grades are often used as indicators of academic achievement but are also a form of teacher ratings (Alvidrez & Weinstein, 1999). They have extensive consequences for students' academic career as they provide feedback for students and parents, are connected with students' academic self-concept, and are in most countries part of the leaving certificate (Südkamp, Kaiser, & Möller, 2012). Notwithstanding the relevance of school grades, future research on effects of teacher-parent congruence on students' academic achievement might use more objective measurements such as student achievement tests.

Finally, our results were based on a cross-sectional study. Hence, conclusions about the causality between teacher-parent congruence in ratings and students' school grades cannot be made. Longitudinal data are needed to examine the (bi-)directionality between ratings and school grades and to explore the intersection of the development of teacher-parent congruence in ratings and students' academic development.

4.5.4 Conclusion

Teacher and parent ratings of students' facets of giftedness matter for students' academic development. Our study offers insights into teacher and parent ratings of teacher-nominated elementary school students by showing similar (low to mediocre) accuracy levels for the two kinds of raters, stronger halo effects for teacher than parent ratings, and low to mediocre congruence in teacher and parent ratings. However, school grades were often best when teachers and parents agreed in high ratings of teacher-nominated students' facets of giftedness. Furthermore, high parent ratings of mathematical abilities and engagement reduced the connection between teacher ratings and students' math grades. Teachers and parents need to be sensitized to the extent of their diagnostic skills to rate students and need to be trained in these skills. Teachers, parents, and others in charge of students' academic development need to consider that teacher and parent ratings can affect students' achievements—separately and conjointly.

5

General Discussion

5 General Discussion

Elementary school teachers' beliefs and judgments about students' giftedness can have far-reaching consequences for students' academic development (e.g., Acar, Sen, & Cayirdag, 2016; Brighton, Moon, Jarvis, & Hockett, 2007; Coleman, Micko, & Cross, 2015), which underscores the relevance of the present dissertation. To gain a comprehensive view, Südkamp, Kaiser, and Möller's (2012) *model of judgment accuracy* was adapted as a heuristic framework. Some relevant factors of this heuristic framework were investigated with three empirical studies. In Section 5.1, the results of the studies will be discussed in accordance with the three topics: teacher beliefs about the nature of giftedness, teacher judgments about students' giftedness, and teacher judgments about facets of giftedness. Some general strengths and limitations of the dissertation will be addressed in Section 5.2. In Section 5.3, implications of the dissertation for future research and, in Section 5.4, implications for educational practice will be discussed. Finally, a conclusion will be given in Section 5.5.

5.1 Discussion of General Findings

5.1.1 Teacher beliefs about the nature of giftedness

The scientific discussion about what giftedness means is far from settled (e.g., Dai, 2009; Mayer, 2005). Also, teachers seem to differ in their beliefs about giftedness (e.g., Copenhaver & McIntyre, 1992; Endepohls-Ulpe & Ruf, 2005; Rubenzer & Twaite, 1979). In Study 1, this dissertation aimed to provide information about how teacher beliefs align with key questions about the meaning of giftedness. Therefore, a questionnaire was developed on the basis of the comprehensive conception of giftedness proposed by Subotnik, Olszewski-Kubilius, and Worrell (2011, 2012, see Section 1.1.2). It assesses eight dimensions of teacher beliefs concerning the content and development of giftedness: *Domain-Specific vs. Holistic Giftedness, Heterogeneity vs. Homogeneity, Importance of Achievement, Importance of Intelligence, Mutable vs. Fixed Giftedness, Interplay of Personal and Environmental Factors, Deliberate Practice, and Different Key Variables for Children and Adults*. The structure of this questionnaire was investigated and differences in beliefs among four different groups of teachers (i.e., elementary school teachers, school teachers who worked in an enrichment program, instructors of the same program who were not school teachers, and student teachers who aimed to become secondary school teachers) were viewed. Furthermore, the relation between beliefs about the nature of giftedness and beliefs about the malleability of intelligence was explored.

In Study 1, the theoretically derived eight dimensions of beliefs about giftedness were empirically supported in the total sample, showing that teachers seem to structurally distinguish between these dimensions. An analysis of teachers' responses on these dimensions indicated that, overall, teachers showed a slight tendency to be in line with the chosen characteristics of Subotnik et al.'s (2011) conception. For example, teachers tended to agree that giftedness is domain-specific and not holistic, that giftedness develops through an interplay of personal and environmental factors and that giftedness means something different for children than for adults. Teachers also tended to agree that intelligence is the most important characteristic of giftedness. Previous research on elementary school teachers also showed the strong connection that teachers see between intelligence and giftedness (Endepohls-Ulpe & Ruf, 2005; Miller, 2009; Schack & Starko, 1990). At the elementary school level, this belief is mainly in alignment with Subotnik et al.'s (2011) conception of giftedness, as they state that the potential to achieve is the key variable at the beginning stage of giftedness.

However, with the exception that teachers rather clearly disagreed with the notion that giftedness is mainly the result of deliberate training—a conception that was, for instance, proposed by Ericsson (2014)—no clear-cut agreement or disagreement was observed. Furthermore, the correlations between the dimensions were low, indicating that teachers did see them as rather independent from each other. Hence, teachers varied strongly in their combinations of dimensions of beliefs. Teachers' rather tentative positioning concerning the nature of giftedness and the observed diversity of beliefs might be due to teachers' uncertainty about the meaning of giftedness in relation to the measured aspects or the belief that nearly everything is possible concerning the conception of giftedness.

Regarding the investigation of the beliefs of the four groups of teachers, three findings need to be highlighted. First, partial strong measurement invariance between groups could be established. All groups differentiated among the eight dimensions, but some items differed in their factor loadings. This result indicates that the groups might have had different understandings of some items, that items might have evoked different associations, or that the groups might have answered on the basis of different response sets (Chen, 2008). Second, student teachers' beliefs differed the most from those of the other three groups. This might be due to their lack of teaching experience in gifted classes, as found in previous research (Endepohls-Ulpe & Ruf, 2005; Schack & Starko, 1990). Student teachers' lack of teaching experience in general classrooms might be another reason, as this variable has been associated with teachers' identification of students as gifted (Guskin Peng, & Simon, 1992; Rubenzer & Twaite, 1979; Siegle, Moore, Mann, & Wilson, 2010), although studies with a concrete focus on beliefs about giftedness mostly found no differences (Baudson & Preckel, 2013a, 2016; Guskin, Peng, & Majd-Jabbari, 1988; Şahin & Düzen, 1994). However, the student teachers in this study were to become secondary school teachers. Therefore, the difference in school level or the possibility that the student teachers thought more about older children might have contributed to the differences. Third, elementary school teachers and the two groups of teachers in the enrichment program did not differ or differed only slightly in their beliefs. This result was unexpected because, as mentioned before, differences in beliefs about giftedness have previously been found to be associated with differences in teaching experience with gifted classes (Endepohls-Ulpe & Ruf, 2005; Schack & Starko, 1990). This result might indicate that the strength of contact with gifted classes or the ratio between general classrooms and gifted classes might be relevant. In this study, teachers and instructors in an enrichment program were viewed and not, for example, teachers of full-time gifted classes.

Teachers' beliefs about the malleability of intelligence were connected to several dimensions of beliefs about giftedness. Teachers were overall rather discordant as to whether intelligence and giftedness are fixed or mutable. But if they saw intelligence as fixed instead of malleable, they tended to see giftedness as fixed instead of changeable. This result indicates that teachers' beliefs about intelligence are integrated, in part, into beliefs about giftedness. Therefore, if beliefs about giftedness are viewed, beliefs about intelligence should be considered, too.

5.1.2 Teacher judgments about students' giftedness

Study 2 investigated teacher judgments about students' giftedness. As intelligence is—particularly at the elementary school level—one of the main facets of most conceptualizations of giftedness, its influence on teacher judgments was examined (see Sternberg & Davidson, 2005). Study 1 showed that elementary school teachers also tended to see intelligence as an integral characteristic of giftedness. Furthermore, students' giftedness has to be seen in relation to other students' abilities; this notion pervades scientific conceptions about giftedness (Freeman, 2005; Ziegler, 2005) and is included in teachers' beliefs about giftedness as well (Lee, 1999; Zhang & Sternberg, 1998). In this study, elementary school teachers' nominations of students for an enrichment program for gifted students were viewed. In addition to considering students' individual level of intelligence, this study focused on the average level of intelligence in a class. Furthermore, teacher variables such as beliefs about giftedness and intelligence and experience in the area of giftedness were included.

In addition to the hypothesis that students with higher intelligence levels would have a higher probability of getting nominated, the hypothesis was formulated that for students with the same intelligence test score, the class-average level of intelligence would be negatively connected to their chances for nomination. These hypotheses were informed by the research literature on reference group effects in teacher-assigned grades (e.g., Südkamp & Möller, 2009) and teacher judgments about cognitive abilities (e.g., Trautwein & Baeriswyl, 2007). The study discriminated between crystallized and fluid intelligence. Furthermore, it explored whether teachers' beliefs in domain-specific or holistic giftedness, beliefs in the mutability of intelligence, and experience in the area of giftedness were related to nominations or moderated the association between the class-average level of intelligence and students' nomination probability.

As expected and in alignment with previous research (Acar et al., 2016; Machts, Kaiser, Schmidt, & Möller, 2016), higher fluid and crystallized intelligence test scores were connected to a higher probability of nomination. The correlations of students' test scores in fluid (r = .39) and crystallized (r = .33) intelligence with teacher nominations were similar to the results of Machts et al.'s (2016) meta-analysis. Importantly, nomination probabilities differed for students with the same fluid or crystallized intelligence scores depending on the class-average level of fluid or crystallized intelligence. Students in classes with lower average levels of intelligence had a higher probability of getting nominated than students in classes with higher average levels of intelligence. Therefore, the study showed that the negative reference group effect of students' abilities on teacher judgments (e.g., Baudson, Fischbach, & Preckel, 2014; Südkamp & Möller, 2009) seems to be generalizable to teacher judgments about students' giftedness. The finding was in contrast to McBee's (2010) finding that more students get identified as gifted if the average level of achievement is high. However, teacher nominations were only one identification method in McBee's study; specifically, one of the other identification methods was achievement tests, and he could not adequately control for students' individual abilities.

In Study 2, the relations of three teacher variables with teachers' nomination decisions and the negative reference group effect were investigated. Teachers' beliefs about the malleability of intelligence were not connected to teacher nominations or the reference group effect. Therefore, although they were related to beliefs about giftedness in Study 1, they did not function as guides for judgments about whether or not a student is gifted. Concerning teachers' experience in the area of giftedness, the results showed that students had a lower probability of getting nominated if their teachers had experience in the area of giftedness. However, when separated by the kind of experience, no type of experience was significantly associated with teachers' nomination decisions, but some kinds of experience strengthened or lessened the negative reference group effect. In comparison to teachers without the following kinds of experiences, teachers who stated that they had attended lectures or courses about giftedness at a university or read literature about giftedness were less prone to the negative reference group effect, and teachers with off-the-job training were more affected by it. On-the-job training and having taught gifted classes were not connected to the reference group effect. However, the reasons for this complex pattern could not be explored in this study and one can only speculate that the results might be due to differences in scientific rigor, foci, or practice orientation among the various types of training.

Teachers' beliefs that giftedness is holistic were connected to the concern that teachers might think that students will make it on their own and do not need special support (Moon, 2009). Therefore, the developed scale about the belief that giftedness is holistic or domain-specific from Study 1 was included in Study 2. Teachers in Study 2 had domain-specific rather than holistic views of giftedness. This belief was not related to students' probability of getting nominated per se but indirectly via a moderation of the negative reference group effect. In classes with low average levels of intelligence, but not in classes with high average levels of intelligence, equally intelligent students' chances of getting nominated were higher if teachers saw giftedness as holistic instead of domain-specific. Therefore, this belief seems to function more as a filter of information than as a direct guide for the nomination decision. Furthermore, focusing on elementary school teachers who nominated students for an extracurricular enrichment program, these results showed that a holistic view of giftedness was not a disadvantage for students. Teachers with more domain-specific views of giftedness might use different indicators for giftedness and might therefore be less affected by the average level of intelligence. For example, Study 1 showed with a small but significant correlation that teachers who saw giftedness as domain-specific tended to see intelligence as less important. Hence, whereas teachers with holistic views might concentrate more on fluid and crystallized intelligence and therefore be more prone to effects of class-average fluid and crystallized intelligence, teachers who see giftedness as domain-specific might focus more on domain-specific abilities and might be, in turn, more sensitive to the class-average levels of these variables.

5.1.3 Teacher judgments of facets of giftedness

Modern conceptions of giftedness see giftedness as a multifaceted construct that includes students' cognitive abilities, creativity, and motivation (see Sternberg & Davidson, 2005). Studies 1 and 2 showed that students' intelligence is important for teachers' beliefs and judgments about students' giftedness. Other research has found that creativity and motivation are also elements of teachers' beliefs about giftedness (e.g. Miller, 2009). Study 3 investigated how teachers judge these facets of giftedness among elementary school students whom the teachers nominated for an enrichment program for gifted students. Specifically, teachers rated the following characteristics: cognitive abilities (i.e., verbal abilities, mathematical abilities, and deductive reasoning), creative thinking, and engagement. Teacher ratings of students' academic achievement but also the congruence between teacher and parent ratings have been found to be related to students' academic achievement (Südkamp et al., 2012; Brenner & Mistry, 2007; Peet, Powell, & O'Donnel, 1997). Therefore, in Study 3, teacher ratings were, first, compared with parent ratings and, second, teacher-parent congruence was set into relation with students' school grades.

As previous research on teacher and parent ratings of facets of giftedness have rarely been based on gifted students (cf. Chan, 2000, for gifted secondary school students), the first objective of this study was to expand the body of research with results on teacher and parent ratings of teacher-nominated gifted elementary school students concerning three topics: (a) whether teacher and parent ratings differed in their accuracy levels, (b) whether teachers and parents were differently affected by halo effects, and, (c) how strongly teachers and parents agreed in their ratings concerning each characteristic.

(a) Teachers' accuracy levels in rating verbal abilities (r = .35), mathematical abilities (r = .26), and deductive reasoning (r = .24) were mediocre at best. Teacher ratings of creative thinking (r = -.05, *ns*) and engagement (r = .15, *ns*) were not significantly correlated with the corresponding student data. Parents' rating accuracy did not differ significantly from teachers' accuracy. Therefore, teachers and parents were equally able judges, with low to mediocre accuracy. For engagement, the results were in line with previous research (Helmke & Schrader, 1989; Spinath 2005); however, for cognitive abilities and creativity (Geiser, Mandelman, Tan, & Grigorenko, 2016; Machts et al., 2016; Miller & Davis, 1992; Sommer, Fink, & Neubauer, 2008), previous research not only indicated higher accuracy levels but also that teachers were more accurate judges than parents. Teachers' and parents' rating accuracy may have been reduced due to students' nomination status. This might have been accompanied by beliefs about how gifted students have to be which did not align with these students' characteristics. Furthermore, the sample in this study only included teacher-nominated gifted students. The variance in the facets of giftedness may have been smaller than in samples from general classrooms, resulting in artificially reduced correlations for statistical reasons (Wild, 1993). Moreover, due to teachers' simultaneous decision to nominate students as gifted, teachers might have (implicitly) used the ratings to confirm their nomination decisions.

(b) As expected and in confirmation of previous work (e.g., Li, Lee, Pfeiffer, & Petscher, 2008; Peters & Gentry, 2010; Petscher & Li, 2008; Urhahne, 2011), teacher and parent ratings were both affected by halo effects, but teacher ratings more strongly than parent ratings. All correlations between teacher ratings ($.49 \leq r \leq .79$) were higher than the correlations between the student data ($|.02| \leq r \leq .27$), indicating that they connected facets of giftedness more strongly than the student data implied. The ratings pointed to rather homogeneous profiles of teacher-nominated gifted students, although the student data did not support such uniformity and additionally showed a high diversity of strengths and weaknesses in students' profiles. Similarly, parent ratings ($.10 \leq r \leq .58$) were more strongly intercorrelated than student data, with the exception of the correlation between rated verbal and rated mathematical abilities. However, the correlations between all teacher ratings were higher than between parent ratings, indicating that teachers differentiated between the different facets of giftedness less than parents did. Fisicaro and Lance (1990) mentioned as one possible reason for halo effects that evaluators might not be able to discriminate between characteristics. The factor analyses in this study did not support this possibility, instead showing the separability of each rated facet of giftedness by teachers and parents. However, halo effects might play a role due to a global impression or a salient characteristic guiding the ratings (Thorndike, 1920). The belief that gifted students are intelligent, creative, and motivated to learn, which has been found for teachers and parents (e.g., Buckley, 1994; Endepohls-Ulpe & Ruf, 2005; Louis, 1992), might have led to the high correlations. Furthermore, Anders, McElvany, and Baumert (2010) investigated teacher ratings and found a global factor for academic achievement that included ratings of cognitive abilities and motivation. As the daily school experiences of teachers are closer to students' academic achievements than parents' experiences, such a factor might have guided the teachers and to a lesser degree the parents in their ratings of facets of giftedness.

One result stood out: Teachers in Studies 1 and 2 stated that they tend to see giftedness as domain-specific. However, in Study 3, teachers nominated students whom they saw as rather similarly strong in verbal and mathematical abilities, although the nominated students differed more strongly in these characteristics as determined through test scores. Therefore, there might be a difference between teachers' explicitly stated beliefs and nomination choices that might reveal their implicit beliefs. The correlation between verbal and mathematical abilities was the lowest of the correlations between teacher ratings. This result might therefore also implicate that within their framework of highly connected characteristics, they had rather domain-specific views. Teachers' understanding of domain-specific might also be that gifted students are rather similarly (good) in several domains but excellent in one or a few.

(c) The correlations between teacher and parent ratings of the same facet were, as expected from previous research (Geiser et al., 2016; Miller & Davis, 1992), medium for verbal (r = .31) and mathematical abilities (r = .40) and weak for creative thinking (r = .10, *ns*). Teacher and parent ratings for deductive reasoning (r = .12, *ns*) were not significantly correlated and both ratings of engagement (r = .28) were rather weakly associated. Both correlations were lower than expected (Chan, 2000; Miller & Davis, 1992). The low to at best mediocre associations between teacher and parent ratings indicated that teachers and parents rated teacher-nominated gifted students very differently overall. These differences might be explained by the different experiences and relationships teachers and parents had with the students, their different opportunities to observe students' characteristics, and the rather high level of inference, low observability, and rather high subjectivity of some items (e.g., "understands abstract ideas", "original solutions").

The second objective of Study 3 was to examine the effects of congruence in teacher and parent ratings for each facet of giftedness on students' German and math grades. In alignment with Brenner and Mistry (2007), teachers and parents' ratings were additively connected to students' German grades. Therefore, students' German grades were best when teachers and parents agreed in their high ratings for each of the facets of giftedness, whereas their German grades were worst when both had congruently low ratings. Concerning math grades, teacher ratings for deductive reasoning and creative thinking, but not parent ratings, were connected to students' math grades. However, for ratings of mathematical abilities and engagement, moderation analyses showed that high parent ratings reduced the association between teacher ratings and math grades. This moderation led, in alignment with Brenner and Mistry's (2007) results, to a situation that when teacher ratings were low, high parent ratings were connected to better math grades than in cases of low parent ratings. Therefore, not teacher ratings alone, but mostly the congruence between teacher and parent ratings, have to be considered in examining students' academic achievement. As the moderation analysis illustrated, teacher-nominated gifted students can benefit from parent ratings of student characteristics when they are higher than teacher ratings, as they soften the effects of teacher ratings on students' school grades. This effect might be due to supportive family variables.

5.2 Strengths and Limitations of the Present Dissertation

Some general strengths and limitations of this dissertation have to be kept in mind when interpreting the findings. The studies in the present dissertation relied on relatively large samples that stemmed from the same state in Germany. In Study 1, a full sample of student teachers from a university course on pedagogical psychology were involved and out of all the teachers in the enrichment program, 80.36% of school teachers and 30.44% of instructors responded. Furthermore, the elementary school teachers from Studies 1 and 2 were from schools that were randomly selected out of a pool of schools that had sent students to the enrichment program in previous years. In Study 3, twelve of 60 academies of the enrichment program participated, and students came from 189 different elementary schools. Although some samples could have been larger, the results might be relatively representative for these groups in this state. However, the results of this dissertation might not be generalizable to, for example, secondary school teachers, other types of gifted education such as acceleration, and other countries.

In this dissertation, participants received no direct explanation for what giftedness is. This approach had the advantage that beliefs and judgments about giftedness could be captured in as unfiltered a way as possible under the chosen methodological approaches. Furthermore, there is—at least at the moment—no conception of giftedness that everyone can agree upon (see Sternberg & Davidson, 2005). However, the teachers and instructors from the enrichment program in Study 1 and the elementary school teachers in Studies 1, 2, and 3 received the following information about the enrichment program for gifted students: that the focus of the project was on STEM subjects and that the general goal of the enrichment program was to cater to the 10% most gifted students in the state of Baden-Württemberg, Germany. This information could have influenced participants' responses in the studies.

In the three studies, variables aligning with most categories of Südkamp et al.'s (2012) *model of judgment accuracy* were studied. Teacher variables such as beliefs about the nature of giftedness and experiences with gifted students were addressed in Studies 1 and 2, student variables like cognitive abilities in Studies 2 and 3, various teacher judgments such as global nominations in Study 2 and judgments based on ratings scales in Study 3. Concerning the characteristics of the criteria, different sources of student data were used, such as tests in Studies 2 and 3 and self-reports in Study 3. Only correlational analyses, not effectivity-efficiency measures, were used in Studies 2 and 3 because of the deliberately undefined giftedness criterion and the drawbacks of effectivity-efficiency measures discussed by, for example, Gagné (1994) and Hoge and Cudmore (1986). Therefore, this dissertation provides a rather comprehensive view on beliefs and judgments about giftedness. The findings concerned three specific topics but each topic could only be touched upon rather briefly. Furthermore, only a few variables were held equal across the three studies, different participants and content areas were examined in the two studies that dealt with teacher judgments (Studies 2 and 3), nor did the three studies build—with the exception of the inclusion of one dimension of beliefs about giftedness from Study 1 in Study 2—on each other's results. Hence, the content of the studies was rarely linked, and open questions resulting from each study were not pursued.

The studies were conducted in the field. Therefore, the external validity of the studies is probably high. However, many factors could not be accounted for and might therefore have unknowingly influenced the results. Specifically, the level of unexplained variance in Studies 2 and 3 indicated a need for further predictors to gain a more comprehensive and realistic understanding of teacher judgments. Furthermore, this dissertation relied strongly on questionnaires and explicit statements of teacher judgments. This approach seems to be appropriate for the goal of measuring the explicit parts of beliefs and analyzing explicit judgments about students. However, insights into implicit parts of beliefs or into the rather informal teacher judgments of students that are frequently conducted in day-to-day school situations are strongly limited. Finally, all studies were cross-sectional. Whereas this design was suitable for the research questions in Studies 1 and 2, the results in Study 3 on the effects of teacher-parent congruence in rating students on students' school grades would have gained explanatory power from the opportunity to analyze causality.

5.3 Implications for Future Research

Diverse implications for future research can be drawn from this dissertation. Some of the most central ones are outlined here. A primary goal of future research should be to replicate the presented results to control for sampling errors and artifacts, to generalize the results, and to verify hypotheses (Schmidt, 2009). For instance, studies that are as close to the presented studies as possible, with new samples drawn from the same population, are needed to investigate whether the presented findings were chance results. Further replication studies should systematically vary aspects such as the constitution of the dependent variables (e.g., a gifted rating scale instead of one global dichotomous nomination decision) to investigate whether other variables were in part responsible for the research findings, or should select participants from different populations (e.g., secondary school teachers, international populations) to explore the generalizability of the findings to larger or other populations.

The newly developed measures need refinement and validation. A new questionnaire of beliefs about the nature of giftedness was developed in Study 1, and new teacher and parent rating scales for facets of giftedness were presented in Study 3. The questionnaire and the rating scales should be revised to further strengthen their measurement quality by, for example, adapting item content and adding new items. Specifically, an investigation is needed to determine what kind of scaling of the rating scales for teacher and parent judgments is appropriate with regard to such aspects as what kind of reference group (e.g., no specification of a reference group, ratings in comparison to same age peers or to the school class) and how many values the scale should have. The factor structure and the reliability of the factors need to be examined with new samples from the same population as well as different populations. Furthermore, measurement invariance analyses are needed to explore the comparability of different groups, and for example, methods such as the think-aloud technique should be used to investigate reasons for the partial measurement invariance of the questionnaire. The content validity should be reviewed, for example, through expert interviews. Similar and more distantly related beliefs about giftedness should be connected with the questionnaire for construct validity. Additionally, the method-dependency of the beliefs measured has to be researched. This could be done, for example, by comparing the presented questionnaire asking for explicit statements with other methods like case vignettes that can be used to measure implicit parts of beliefs. Also, the associations of each rating scale for a facet of giftedness with other rating scales and student data for (un)related constructs should be investigated for construct and criterion validity.

There is a need for theoretical models of teacher beliefs and judgments about students' giftedness. Although a relatively broad body of research on beliefs and judgments about giftedness exists (e.g., Acar et al., 2016; Baudson & Preckel, 2016; Endepohls-Ulpe & Ruf, 2005; Hany, 1997; Machts et al., 2016; Miller, 2009), there are no theoretical models that systemize these beliefs and judgments across the existing literature and specify which factors are associated with them (and how). Concerning beliefs about giftedness, Study 1 proposed the use of a conception of giftedness as a starting point to extract relevant dimensions of beliefs and to systemize them into two categories: beliefs about the content of giftedness and beliefs about its development. This approach was constructive for the purpose of Study 1 and might be a starting point for similar endeavors. Systematizations and content analyses of conceptions of giftedness can be used to specify a reasonable set of facets of giftedness, of possible connections between facets, and of developmental and educational aspects. However, several further areas of beliefs about giftedness have already been investigated, such as Sternberg and Zhang's

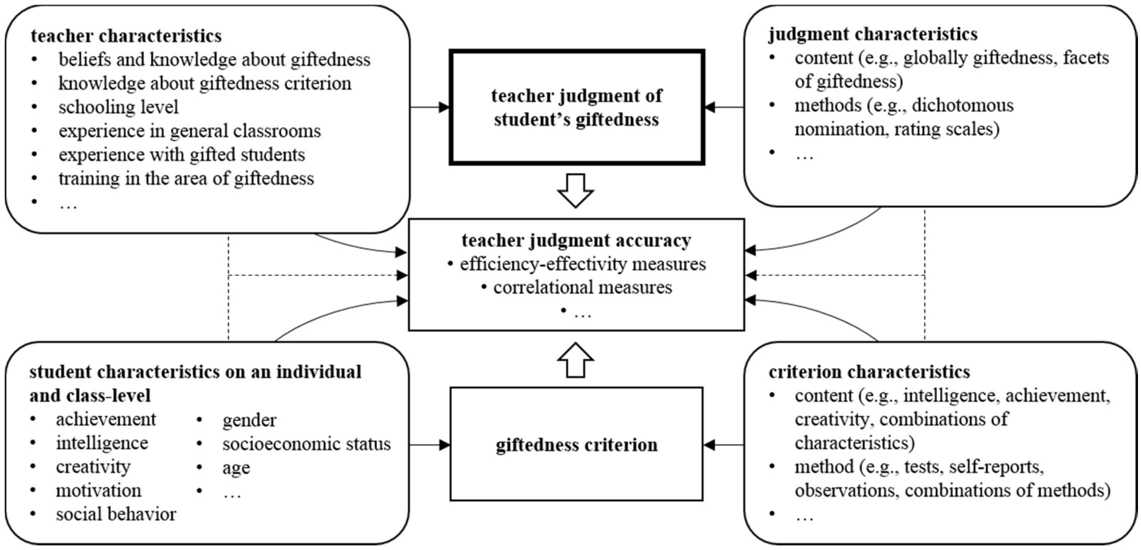

Figure 5.1. Heuristic framework with factors that can contribute to teacher judgments about students' giftedness, adapted from Südkamp et al. (2012).

(1995) pentagonal theory of giftedness; the collections of believed indicators of giftedness by, for instance, Busse, Dahme, Wagner, and Wieczerkowski (1986) and Miller (2009); and the harmony-disharmony-hypotheses by Baudson (2016; Baudson & Preckel, 2013a). These should be connected and integrated into a model about the system of beliefs about giftedness so that predictions about (directions of) dependencies and functions of beliefs can be made and researched.

The situation is similar for teacher judgments about students' giftedness, as there is no model here either. Südkamp et al.'s (2012) *model of judgment accuracy* might be a suitable framework for systemizing influences on teacher judgments about giftedness (see Figure 5.1). However, the model's strength lies mainly in the systemization of factors, less in the deduction of predictions about how judgments are made. Specifically, Brunswik's (1955; see also Hartwig & Bond, 2011 for an application) idea to use probabilities to express the strengths of associations might be useful for a model of teacher judgments about students' giftedness. He proposed probabilities to indicate how strongly proximal (i.e., observable) cues are related to the distal (i.e., not observable) characteristic that should be judged, and probabilities for how strongly proximal cues are related to teachers' judgment. The probability of teachers using certain proximal characteristics such as students' language style, quantity and quality of solving tasks, and handling of homework could be modeled and compared to the probabilities that the proximal cues are indicators for students' giftedness. Additionally, different factors' probabilities of influencing the relationships of proximal cues with the distal characteristic (e.g., family background) and the teacher judgment (e.g., reference group effects) need to be added. To sum up, Südkamp et al.'s (2012) model might be an adequate model for teacher judgments of students' giftedness but needs to be expanded to include indicators of (conditional) strengths and directions of relations between factors and judgments that are derived from research or theory. This would help specify the ways in which different factors such as teacher beliefs and reference group effects affect teacher judgments, help make predictions, and guide further research.

The three studies of this dissertation suggest several possibilities for combinations of the researched content. Five topics for combinations are highlighted in the following paragraphs. First, future studies should use direct and indirect measures of beliefs about giftedness to follow up on the result that a direct measurement via questionnaire indicated rather domain-specific views on giftedness, whereas an indirect measurement via a comparison of teacher judgments of different facets of giftedness among teacher-nominated gifted students could be interpreted as pointing to a rather holistic view of giftedness. For instance, future research might combine both measurements in field studies on teacher judgments of giftedness to explore the relation between direct and indirect measurements. Or in a more experimental approach, investigations of teacher judgments about giftedness might be based on case vignettes that describe students with differently strong homogenous and heterogeneous profiles on different ability levels. These vignettes should be supplemented with questionnaires or interviews about teachers' beliefs about giftedness and their judgment choices.

Second, future research should not only investigate the influence of beliefs about giftedness on judgments about (facets of) giftedness but also their influence on other variables' associations with judgments. Study 2 showed that teachers' beliefs about whether giftedness is holistic or domain-specific had no direct effect on teachers' nomination decisions, but did have an indirect one via a moderation of the reference group effect. Therefore, for instance, in addition to setting teacher ratings of student characteristics into relation to teacher judgments about students' giftedness, as has been done by Neber (2004), it should be measured how strongly which characteristic is seen as indicative of giftedness by teachers. This information should be set into relation with teacher judgments but also with the associations between student characteristics on an individual and class level and teacher judgments. With this approach, different functions of beliefs about giftedness might be localized and a clearer picture of judgment formations might appear.

Third, future research should aim to combine global teacher judgments about students' giftedness with teacher judgments of facets of giftedness. Additionally, the connection between student characteristics and these kinds of judgments should be investigated: for example, how students' intelligence, creativity, and motivation are associated with judgments of intelligence, creativity, motivation, and giftedness. Burke, Haworth, and Ware (1982) indicated that ratings of learning and academic abilities seemed to be the separators for judgments about giftedness rather than, for example, creativity. However, they did not have student data for the rated characteristics to indicate whether the student characteristics themselves were also separators. Urhahne (2011) had teacher ratings and student data about cognitive abilities, creativity, and motivation, but used a giftedness criterion defined post-hoc rather than an actual teacher judgment about giftedness to form conclusions about the accuracy of the different judgments. Hence, an integration is needed to compare, for instance, how student data and teacher judgments of facets of giftedness align with global judgments about students' giftedness, to set the accuracy levels of the judgments about facets of giftedness into relation with the accuracy level of the global judgment, and to investigate how halo effects on teacher judgments of facets differ in relation to the judgment whether or not a student is gifted.

Fourth, there is a need for more detailed measurement of the kind and content of experiences teachers have in the area of giftedness. Results from the present dissertation showed that teaching experience in gifted classes resulted in only a few associations with beliefs about giftedness and none with teacher judgments about giftedness. Some other studies, however, found connections with teachers' perceptions of gifted students (e.g., Guskin et al., 1992; Rubenzer & Twaite, 1979). Also, in this dissertation, a complex pattern of indirect effects of different kinds of trainings in the area of giftedness on teacher nominations emerged. Although some studies have described their trainings to some extent (e.g., Gear, 1978; Goodnough, 2000; Şahin & Cetinkaya, 2015), studies that investigate what exactly teachers know or have learned about giftedness or what experiences teachers have had with gifted students before setting teacher experience into relation with beliefs or judgments about giftedness are rare. Hence, it is hardly possible to infer why different kinds of experiences did or did not have certain effects across the studies. Future research should aim to measure knowledge and beliefs about giftedness before and after training, single out certain aspects of trainings for in-depth investigation, or concentrate on specific kinds of teacher experiences with gifted students.

Fifth, Study 3 stressed the effect that teacher-parent congruence in ratings of teacher-nominated gifted students' characteristics can have on their school grades. Hence, further inquiries should be conducted to follow up on why teacher-parent congruence in judgments of facets of giftedness was mediocre at best. An exploration and comparison of the situations that teachers and parents see as indicators for the evaluated characteristics, of their beliefs about giftedness, and of their judgments regarding whether or not a student is gifted might help to determine the reasons for this result. Furthermore, the (bi)directionality between both ratings and between ratings and school grades should be investigated.

Finally, this dissertation is based solely on cross-sectional data. Some results should be followed up with longitudinal research. For example, beliefs are often found to be rather stable (Baudson & Preckel, 2013a, 2016; Pajares, 1992), although some studies have found changes in beliefs (e.g., Miller, 2009; Schack & Starko, 1990). The beliefs measured in this dissertation did not differ at all according to teachers' numbers of years of experience in general classrooms and differed only slightly for the number of years teachers spent teaching in gifted classes. As changes within teachers were concealed by the chosen methodological approach, the results should be reviewed in a longitudinal design. Moreover, a longitudinal design should be used for the already mentioned follow-up to Study 3 examining the directionalities between teacher and parent ratings and school grades.

5.4 Educational Implications

The present dissertation provides important information which could sensitize teachers, trainers in the area of giftedness, and managers of gifted education programs. Furthermore, it presents a questionnaire about beliefs concerning giftedness and rating scales for facets of giftedness that—following further revisions and validations—might prove to be useful in praxis.

In line with Hoge and Cudmore (1986), teachers should be supported in their task of nominating students as gifted by means of adequate information and preparation and by way of appropriate tools for their judgments. Furthermore, their judgments should be used in combination with other sources such as tests or work samples. The results of

Studies 2 and 3, that teacher judgments were related to reference group effects and halo effects and had low to mediocre accuracy levels, underpin these demands. Previous studies have shown that teacher judgments about giftedness can be modified by trainings (Gear, 1978; Şahin & Cetinkaya, 2015)—however, not always (McCoach & Siegle, 2007; Miller, 2009). Furthermore, Study 2 indicated different effects for different kinds of teacher trainings. Effective trainings connect the transfer of content knowledge with an improvement in assessment knowledge in the content area (Lipowsky & Rzejak, 2012). In their model of competence stages for the identification and support of gifted students, Heller, Reimann, and Senfter (2005) stressed that elementary school teachers need to know, among other things, about scientific conceptions and beliefs about giftedness and about issues surrounding the identification of gifted students. Furthermore, training in the different areas of the model should be flexibly based on teachers' previous knowledge and competencies.

In Study 1, teachers could distinguish between different key characteristics of giftedness, and their beliefs tended to be in accordance with modern conceptions of giftedness. However, clear agreement or disagreement with the notions about giftedness was rare and a rather high diversity of beliefs was observed. An in-depth discussion of the content and development of giftedness might therefore be beneficial for teachers' understanding of giftedness. Furthermore, teachers' beliefs about giftedness can be relevant for the transfer of knowledge, as beliefs can filter information, guide tasks, and form behavior (Five & Bühl, 2012). Although beliefs are often rather stable (Pajares, 1992), beliefs about giftedness themselves have been targets of attempted modification (e.g., Goodnough, 2000; Miller, 2009). Hence, a diagnosis of teachers' beliefs about giftedness can be useful. The presented questionnaire from Study 1 might serve (as an orientation) for such an assessment. At the beginning of a training program, it can help in planning the course of action. An assessment can be useful to structure and guide discussions about the meaning of giftedness and can sensitize trainers to possible reasons for agreement or disagreement to certain notions surrounding the construct. For the participating teachers, it might support their awareness of their own beliefs and help make them amenable to discussion. Moreover, a reassessment of teachers' beliefs after the training can be used to evaluate training effects.

Focusing specifically on assessment knowledge, the *Kultusministerkonferenz* (2015) recently claimed that teachers should be able to identify students who show excellence or have the potential for high achievement. They stressed the importance of training to improve teachers' diagnostic skills. For instance, Helmke (2012, based on Wahl, Huber, & Weinert, 2006) proposed the following training cycle. Teachers should select a student characteristic such as giftedness for which they want to review their diagnostic skills. In the next two steps, teachers evaluate students and these students get tested. Then, teacher judgments and student data are compared using, for example, difference scores or correlations. In the last step before the circle might start again, reasons for discrepancies are discussed. This approach could be based on the rating scales from Study 3 or similar rating scales. In the beginning, reasons for the inclusion of the chosen characteristics to measure giftedness (i.e., cognitive ability, creative thinking, and engagement) can be worked out, debated, and supplemented with further characteristics. Additionally, the included items can be discussed vis-à-vis their value as indicators for the rated characteristics, and for their observability, level of inference, and subjectivity. They might be adapted and complemented by further items. Different approaches to combine the ratings into a judgment about giftedness can be constructed and tested. In this way, teachers can be sensitized to the methodological challenges of the tools that support teacher judgments.

Concerning the collection of student data, it might prove difficult to test students; hence, alternatives such as case vignettes about fictitious students with different features and different behavior might be created. For instance, Fiedler, Walther, Freytag, and Plessner (2002) presented the *Simulated Classroom* that simulates classroom situations in which student answers and characteristics (e.g., for achievement and engagement, Kaiser, Retelsdorf, Südkamp, & Möller, 2013) can be experimentally manipulated. This approach reduces the complexity of the situation and can be designed prototypically to discuss specific issues such as underachievement, but, of course, will be partially unrealistic. Regardless of the source of student data, methodological reasons should be considered when the accuracy or rather the discrepancies between judgments and measurements of student data are debated. The challenges of measuring student characteristics such as giftedness (e.g., ceiling effects of tests; Zettler, Thoemmes, Hasselhorn, & Trautwein, 2014) and creativity (e.g., problems operationalizing creativity, Runco, 1993) have to be addressed. Furthermore, teachers need to be aware of the problems involving self-reports. For instance, self-reports and observable behavior concerning a characteristic can differ (Jamieson-Noel & Winne, 2003).

This dissertation can have practical implications for those who are responsible for gifted education programs. For instance, a self-reflection on their own beliefs about the facets and development of giftedness can be beneficial for their task of deciding who they are looking for. Sternberg, Jarvin, and Grigorenko (2011) proposed four stages of sophistication in the selection of a giftedness definition for a gifted education program. At the lowest stage, a gifted definition is used because it has been used before. There is no self-reflection on the meaning of giftedness. At the next stage, several conceptions of giftedness are viewed and one of them is selected and defended on the basis of the theorists' authority. At Stage 3, program managers can argue why they might use identification methods that do not fit any particular conception of giftedness. At Stage 4, program managers reflect on what they believe giftedness is and whom they want to support. On the basis of this, they select a conception or a combination of conceptions.

The dissertation has shown that just providing teachers with rating scales does not seem to lead to teacher judgments with high accuracy. Furthermore, due to halo effects, teacher judgments might not yield rated student profiles that adequately reflect students' strengths and weaknesses. Therefore, they might provide gifted education programs with only limited information about the qualifications their participants bring with them. Also, the result of reference group effects on teacher judgments of giftedness should receive attention from program managers. They should be aware that teachers orient their judgments on the class—this applies not only to giftedness judgments but to other judgments as well (e.g., Baudson et al., 2014; Trautwein & Baeriswyl, 2007). This tendency can be contrary to the program's stated standards (e.g., to nominate the 5% most gifted students in a state), and can lead to a group of participants that vary (greatly) in their ability levels. Moreover, students from the target group might be overlooked. In this dissertation, the reference group effect of class-average intelligence differed in relation to teachers' beliefs about whether giftedness is domain-specific or holistic and with respect to several kinds of experiences in the area of giftedness. However, it is unknown whether teachers who were not as strongly prone to the class-average level of intelligence were affected by the class or school composition with regard to other characteristics. Therefore, whether and how reference group effects on teacher judgments can be eliminated or changed has not been settled. In conclusion, even though gifted education programs should train teachers and provide them with the tools to identify students in accordance with the program's specifications, program managers should consider combining teacher nominations with other sources to reduce the impact of reference group effects, to increase the fit between student characteristics and the program's specifications, and to supplement further information about students' strengths and weaknesses.

Finally, previous research (for a review, see Glueck & Brenner, 2014) has stressed the importance of creating constructive school-family collaborations with joint responsibility for students' development. These collaborations are effective when teachers and parents are, for example, congruent in their types of support, values, and standards (Christenson & Peterson, 1998). In addition, the present dissertation showed for teacher-nominated gifted students that teachers' and parents' ratings of facets of giftedness were related to students' German grades in an additive manner. Furthermore, high parent ratings of mathematical ability and engagement reduced the connection between teacher ratings and students' math grades. Teachers and parents need to be sensitized to these effects. Specifically, as teacher and parent ratings overall correlated to a low to mediocre degree, it might be advisable for teachers and parents to communicate about their judgments of students, discuss reasons for differences, and focus on students' strengths.

5.5 Conclusion

This dissertation provided in-depth information about elementary school teachers' beliefs and judgments about giftedness and started to close research gaps. Three empirical studies were conducted showing, first, teachers' beliefs concerning the content and development of giftedness. Student teachers' beliefs were markedly different from those of the other groups of teachers. Elementary school teachers and the two groups of teachers in an enrichment program for gifted elementary school students were on average rather similar to each other. However, teachers overall expressed a huge variety of beliefs. Second, negative reference group effects of the class average level of intelligence on students' probability of getting nominated as gifted by teachers were found. These effects were partially influenced by teacher variables. Third, concerning teacher-nominated gifted students, a comparison of teacher and parent ratings of their facets of giftedness showed similar (low to mediocre) accuracy levels, stronger halo effects for teacher ratings than parent ratings, and low to mediocre congruence in teacher and parent ratings. Students' German grades were best when teachers and parents agreed in their high ratings.

The connection between teacher ratings of mathematical abilities and engagement and students' math grades were reduced when parents had high ratings. This dissertation has provided new starting points for research on beliefs about giftedness, reference group effects, and teacher judgments about giftedness and facets of giftedness. The results can be informative and beneficial to teachers and parents as well as trainers and managers of gifted education programs.

References

Acar, S., Sen, S., & Cayirdag, N. (2016). Consistency of the performance and nonperformance methods in gifted identification: A multilevel meta-analytic review. *Gifted Child Quarterly, 60*, 81–101. http://doi.org/10.1177/0016986216634438

Alvidrez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology, 91*, 731–746. http://doi.org/10.1037/ 0022-0663.91.4.731

Ambrose, D. (2002). Socioeconomic stratification and its influences on talent development: Some interdisciplinary perspectives. *Gifted Child Quarterly*, *46*, 170–180. http://doi.org/ 10.1177/001698620204600302

Anastasiow, N. J. (1964a). Frame of reference of teachers' judgments: The psycho-physical model applied to education. *Psychology in the Schools, 1*, 392–395. http://doi.org/10.1002/1520-6807(196410)1:4<392::AID-PITS2310010411>3.0.CO;2-X

Anastasiow, N. J. (1964b). Maximizing identification of the gifted. *The Journal of Educational Research, 57*, 538–541. http://doi.org/10.1177/001698627602000416

Anders, Y., McElvany, N., & Baumert, J. (2010). *Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte?* [Assessment of learning-relevant student characteristics at the transition from primary to secondary school: How differentiated are teachers' judgments?]. In K. Maaz, J. Baumert, C. Gresch, & N. McElvany (Eds.), *Der Übergang von der Grundschule in die weiterführende Schule* (pp. 313-330). Bonn, Germany: Federal Ministry for Education and Research.

Ashman, S. S., & Vukelich, C. (1983). The effect of different types of nomination forms on teachers' identification of gifted children. *Psychology in the Schools, 20*, 518–527. http://doi.org/10.1002/1520-6807(198310)20:4<518::AID-PITS2310200421>3.0.CO;2-B

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 397–438. http://doi.org/10.1080/ 10705510903008204

Babad, E., Bernieri, F., & Rosenthal, R. (1989). Nonverbal communication and leakage in the behavior of biased and unbiased teachers. *Journal of Personality and Social Psychology, 56, 89–94*. http://doi.org/10.1037/0022-3514.56.1.89

Babad, E. Y., Inbar, J., & Rosenthal, R. (1982). Teachers' judgment of students' potential as a function of teachers' susceptibility to biasing information. *Journal of Personality and Social Psychology, 42*, 541. http://doi.org/10.1037/0022-3514.42.3.541

Baeriswyl, F., Wandeler, C., & Trautwein, U. (2011). „Auf einer anderen Schule oder bei einer anderen Lehrkraft hätte es für's Gymnasium gereicht!" Eine Untersuchung zur Bedeutung von Schulen und Lehrkräften für die Übertrittsempfehlung [Another school or another teacher would have recommended a higher track: The meaning of schools and teachers for secondary tracking recommendations]. *Zeitschrift Für Pädagogische Psychologie [German Journal of Educational Psychology], 25*, 39–47. doi:10.1024/1010-0652/ a000028

Bain, S. K., Choate, S. M., & Bliss, S. L. (2006). Perceptions of developmental, social, and emotional issues in giftedness: Are they realistic? *Roeper Review, 29*, 41–48.

Bangel, N. J., Moon, S. M., & Capobianco, B. M. (2010). Preservice teachers' perceptions and experiences in a gifted education training model. *Gifted Child Quarterly*, *54*, 209–221. http://doi.org/10.1177/0016986210369257

Baudson, T. G. (2016). The mad genius stereotype: Still alive and *well. Frontiers in Psychology, 7, 1–9*. http://doi.org/10.3389/fpsyg.2016.00368

Baudson, T. G., Fischbach, A., & Preckel, F. (2014). Teacher judgments as measures of children's cognitive ability: A multilevel analysis. *Learning and Individual Differences*. http://doi.org/10.1016/j.lindif.2014.06.001

Baudson, T. G., & Preckel, F. (2013b). Development and validation of the German Test for (Highly) Intelligent Kids—T (H) INK. *European Journal of Psychological Assessment, 29*, 171–181. http://doi.org/10.1027/1015-5759/a000142

Baudson, T. G., & Preckel, F. (2013a). Teachers' implicit personality theories about the gifted: An experimental approach. *School Psychology Quarterly, 28*, 37–46. http://doi.org/10.1037/spq0000011

Baudson, T. G., & Preckel, F. (2016). Teachers' conceptions of gifted and average-ability students on achievement-relevant dimensions. *Gifted Child Quarterly, 60*, 212 – 225. http://doi.org/10.1177/0016986216647115

Bégin, J., & Gagné, F. (1994). Predictors of attitudes toward gifted education: A review of the literature and a blueprint for future research. *Journal for the Education of the Gifted*, *17*, 161–179. http://doi.org/10.1177/016235329401700206

Bégin, J., & Gagné, F. (1995). Predictors of a general attitude toward gifted education. *Journal for the Education of the Gifted, 18*, 74–86. http://doi.org/10.1177/016235329401800106

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*, 289–300. http://doi.org/10.2307/2346101

Berlin, J. E. (2009). It's all a matter of perspective: Student perceptions on the impact of being labeled gifted and talented. *Roeper Review, 31*, 217–223. http://doi.org/10.1080/027831 90903177580

Bianco, M., Harris, B., Garrison-Wade, D., & Leech, N. (2011). Gifted girls: Gender bias in gifted referrals. *Roeper Review, 33*, 170–181. http://doi.org/10.1080/02783193.2011.580500

Bianco, M., & Leech, N. L. (2010). Twice-exceptional learners: Effects of teacher preparation and disability labels on gifted referrals. *Teacher Education and Special Education, 33*, 319–334. http://doi.org/10.1177/0888406409356392

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246–263. http://doi.org/10.1111/j.1467-8624.2007.00995.x

Borland, J. (1978). Teacher identification of the gifted: A new look. *Journal for the Education of the Gifted, 2*, 22–32.

Borland, J. H. (1997). The construct of giftedness. *Peabody Journal of Education, 72*, 6–20.

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* (2nd ed., pp. 1-19). New York: Cambridge University Press.

Borland, J. H. (2009). Myth 2: The gifted constitute 3% to 5% of the population. Moreover, giftedness equals high IQ, which is a stable measure of aptitude: Spinal tap psychometrics in gifted education. *Gifted Child Quarterly, 53*, 236–238. http://doi.org/10.1177/0016986 209346825

Brenner, A. D., & Mistry, R. S. (2007). Congruence of mother and teacher educational expectations and low-income youth's academic competence. *Journal of Educational Psychology, 99*, 140. http://doi.org/ 10.1037/0022-0663.99.1.140

Brighton, C. M., Hertberg, H. L., Moon, T. R., Tomlinson, C. A., & Callahan, C. M. (2005). *The feasibility of high-end learning in a diverse middle school*. Storrs, CT: National Research Center on the Gifted and Talented.

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*. Connecticut, US: National Research Center on the Gifted and Talented.

Brody, L., & Stanley, J. (2005). Youths who reason exceptionally well mathematically and or verbally. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 20-37). New York: Cambridge University Press.

Brown, S. W., Renzulli, J. S., Gubbins, E. J., Siegle, D., Zhang, W., & Chen, C.-H. (2005). Assumptions underlying the identification of gifted and talented students. *Gifted Child Quarterly, 49*, 68–79. http://doi.org/10.1177/001698620504900107

Brunner, M., Anders, Y., Hachfeld, A., & Krauss, S. (2013). The diagnostic skills of mathematics teachers. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 229–248). Springer. http://doi.org/10. 1007/978-1-4614-5149-5_11

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217. http://doi.org/10.1037/h0047470

Buckley, K. C. (1994). Parent's views on education for the gifted. *Roeper Review, 16*, 215–216. http://doi.org/10.1080/02783199409553578

Buhi, E. R., Goodson, P., & Neilands, T. B. (2008). Out of sight, not out of mind: Strategies for handling missing data. *American Journal of Health Behavior, 32*, 83–92. http://doi.org/10.5993/AJHB.32.1.8

Burke, J. P., Haworth, C. E., & Ware, W. B. (1982). Scale for rating behavioral characteristics of superior students: An investigation of factor structure. *The Journal of Special Education, 16*, 477–485. http://doi.org/10.1177/0022466982016 00411

Busse, T. V., Dahme, G., Wagner, H., & Wieczerkowski, W. (1986). Teacher perceptions of highly gifted students in the United States and West Germany. *Gifted Child Quarterly, 30*, 55–60. http://doi.org/10.1177/001698628603000202

Callahan, C. M. (2005). Identifying gifted students from underrepresented populations. *Theory Into Practice*, *44*, 98–104. http://doi.org/10.1207/s15430421tip4402_4

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

Carr, P. B., Rattan, A., & Dweck, C. S. (2012). Implicit theories shape intergroup relations. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology* (Vol. 45, pp. 127-165). Burlington: Academic Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Cashion, M., & Sullenger, K. (2000). “Contact us next year”: Tracing teachers’ use of gifted practices. *Roeper Review, 23*, 18–21. http://doi.org/10.1080/0278319000955 4056

Chan, D. W. (2000). Exploring identification procedures of gifted students by teacher ratings: Parent ratings and student self-reports in Hong Kong. *High Ability Studies, 11*, 69–82. http://doi.org/10.1080/713669176

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. http://doi.org/10.1080/10 705510701301834

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*, 1005–1018. http://doi.org/10.1037/ a0013193

Chessman, A. (2010). *Teacher attitudes and effective teaching practices for gifted students at Stage 6* (Doctoral dissertation). Retrieved from http://handle.unsw. edu.au/1959.4 /45448

Christenson, S. L. (1999). Families and schools: Rights, responsibilities, resources, and relationships. In R. C. Pianta & M. J. Cox (Eds.), *The transition to kindergarten* (pp. 143–177). Baltimore, MD: ERIC.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Coleman, L. J., & Cross, T. L. (1988). Is being gifted a social handicap? *Journal for the Education of the Gifted, 11*, 41–56. http://doi.org/10.1177/0162353214521486

Coleman, M. R., & Gallagher, J. J. (1995). State identification policies: Gifted students from special populations. *Roeper Review*, *17*, 268–275. http://doi.org/10.1080/02783199509553681

Coleman, L. J., Micko, K. J., & Cross, T. L. (2015). Twenty-five years of research on the lived experience of being gifted in school. Capturing the students' voices. *Journal for the Education of the Gifted, 38*, 358–376. http://doi.org/10.1177/0162353215607322

Copenhaver, R. W., & McIntyre, D. J. (1992). Teachers' perception of gifted students. *Roeper Review*, *14*, 151–153. http://doi.org/10.1080/02783199209553411

Côté, J. (1999). The influence of the family in the development of talent in sport. *The Sport Psychologist, 13*, 395–417.

Cramond, B., & Martin, C. E. (1987). Inservice and preservice teachers' attitudes toward the academically brilliant. *Gifted Child Quarterly*, *31*, 15–19. http://doi.org/10.1177/001698 628703100103

Cross, T. L. (2005). Nerds and geeks: Society's evolving stereotypes of our students with gifts and talents. *Gifted Child Today, 28*, 26–28.

Cross, T. L., & Coleman, L. J. (2005). School-based conception of giftedness. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 52–63). New York, US: Cambridge University Press.

Cudeck, R. (2000). Exploratory factor analysis. In H. E.A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 265–296), San Diego, California: Academic Press. http://doi.org/10.1016/B978-012691360-6/ 50011-2

Dai, D. Y. (2009). Essential tensions surrounding the concept of giftedness. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 39–80). Springer. http://doi.org/10.1007/ 978-1-4020-6162-2 3

Davidson, J. E. (2009). Contemporary models of giftedness. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 81–97). Springer Netherlands. http://doi.org/ 10.1007/978-1-4020-6162-2 4

de Boer, H., Bosker, R. J., & van der Werf, M. P. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology, 102*, 168. http://doi.org/10.1037/a0017289

Deku, P. (2013). Teacher nomination of gifted and talented children: A study of basic and senior high schools in the central region of Ghana. *Journal of Education and Practice, 4*, 1–7.

Dicke, A.-L., Lüdtke, O., Trautwein, U., Nagy, G., & Nagy, N. (2012). Judging students' achievement goal orientations: Are teacher ratings accurate? *Learning and Individual Differences, 22*, 844–849. doi:10.1016/j.lindif.2012.04.004

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198. http://doi.org/10.1037/0022-0663.98.1.198

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*, 267–285. http://doi.org/10.1207/s15327965pli0604_1

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, US: Psychology Press.

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry, 6*, 267–285. http://doi.org/10.1207/s15327965pli0604_1

Elhoweris, H., Mutua, K., Alsheikh, N., & Holloway, P. (2005). Effect of children's ethnicity on teachers' referral and recommendation decisions in gifted and talented programs. *Remedial and Special Education, 26*, 25–31. http://doi.org/10.1177/074 19325050260010401

Endepohls-Ulpe, M. (2008). Gender, social behaviour and domain of ability. Influences on teachers' diagnoses of giftedness. *Gifted and Talented International, 23*, 125–33.

Endepohls-Ulpe, M., & Ruf, H. (2005). Primary school teachers' criteria for the identification of gifted pupils. *High Ability Studies*, *16*, 219–228. http://doi.org/10.1080/13598130600618140

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*, 128–141. http://doi.org/10.1207/S15328007SEM0801_7

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Ericsson, K. A. (2014). Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms. *Intelligence, 45*, 81–103. http://doi.org/10.1016/j.intell.2013.12.001

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2005). Giftedness viewed from the expert-performance perspective. *Journal for the Education of the Gifted, 28*, 287–311. http://doi.org/10.4219/jeg-2005-335

Eyre, D., Coates, D., Fizpatrick, M., Higgins, C., McClure, L., Wilson, H., & Chamberlin, R. (2002). Effective teaching of able pupils in the primary school: The findings of the Oxfordshire effective teachers of able pupils project. *Gifted Education International, 16*, 158–169. http://doi.org/10.1177/026142940201600210

Fatouros, C. (1986). Early identification of gifted children is crucial… But how should we go about it? *Gifted Education International, 4*, 24–28. http://doi.org/10.1177/026142948600400107

Feldhusen, J. F. (1986). A conception of giftedness. In K. A. Heller & J. F. Feldhusen (Eds.), *Identifying and nurturing the gifted. An international perspective* (pp. 33-38), Stuttgart: Hans Huber Publishers.

Feldhusen, J. F. (2005). Giftedness, talent, expertise, and creative achievement. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 64–79), New York: Cambridge University Press.

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics, 21*, 55–69. http://doi.org/10.1007/BF00311015

Fiedler, K., Walther, E., Freytag, P., & Plessner, H. (2002). Judgment biases in a simulated classroom—A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes, 88*, 527–561. http://doi.org/10.1006/obhd.2001.2981

Fischbach, A., Baudson, T. G., Preckel, F., Martin, R., & Brunner, M. (2013). Do teacher judgments of student intelligence predict life outcomes? *Learning and Individual Differences, 27*, 109–119. http://doi.org/10.1016/j.lindif.2013.07.004

Fischer, C. (2014). *Individuelle Förderung als schulische Herausforderung* [Individual support as challenge for school]. Berlin: Friedrich-Ebert-Stiftung.

Fischer, C., & Müller, K. (2014). Gifted education and talent support in Germany. *CEPS Journal, 4*(3), 31.

Fisicaro, S. A., & Lance, C. E. (1990). Implications of three causal models for the measurement of halo error. *Applied Psychological Measurement, 14*, 419–429. http://doi.org/10.1177/014662169001400407

Fives, H., & Buehl, M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook: Volume 2 Individual differences and cultural and contextual factors* (p. 471–499). Washington: American Psychological Association.

Ford, D. Y. (1998). The underrepresentation of minority students in gifted education problems and promises in recruitment and retention. *The Journal of Special Education*, *32*, 4–14. http://doi.org/10.1177/002246699803200102

Foreman, J. L., & Gubbins, E. J. (2014). Teachers see what ability scores cannot: Predicting student performance with challenging mathematics. *Journal of Advanced Academics*, 1–19. http://doi.org/10.1177/1932202X14552279

Frasier, M. M., & Passow, A. H. (1994). *Towards a new paradigm for identifying talent potential* (Research Monograph 94112). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.

Freeman, J. (2005). Permission to be gifted: How conceptions of giftedness can change lives. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 80–97). New York: Cambridge University Press.

Freeman, J., & Josepsson, B. (2002). A gifted programme in Iceland and its effects. *High Ability Studies*, 13, 35–46. http://doi.org/10.1080/13598130220132299

Friedman-Nimz, R. (2009). Myth 6: Cosmetic use of multiple selection criteria. *Gifted Child Quarterly, 53*, 248–250. http://doi.org/10.1177/0016986209346925

Gagné, F. (1994). Are teachers really poor talent detectors? Comments on Pegnato and Birch's (1959) study of the effectiveness and efficiency of various identification techniques. *Gifted Child Quarterly*, *38*, 124–126. http://doi.org/10.1177/00169862 9403800305

Gagné, F. (1998). A proposal for subcategories within gifted or talented populations. *Gifted Child Quarterly, 42*, 87–95. http://doi.org/10.1177/001698629804200203

Gagné, F. (2005). From gifts to talents: The DMGT as a developmental model. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 98–119). New York: Cambridge University Press.

Gagné, F., & St Père, F. (2001). When IQ is controlled, does motivation still predict achievement? *Intelligence, 30*, 71–100. http://doi.org/10.1016/S0160-2896(01) 00068-X

Ganzeboom H. B. (2010, May). *A new International Socio- Economic Index (ISEI) of occupational status for the International Standard Classification of Occupations 2008 (ISCO-08) constructed with data from the ISSP 2002–2007*. Paper presented at the annual meeting of the International Social Survey Programme, Lisbon, Portugal.

Ganzeboom, H. B., & Treiman, D J. (2011). *International stratification and mobility file: Conversion tools.* Amsterdam: Department of Social Research Methodology, Retrieved from http://www.harryganzeboom.nl/ismf/index.htm.

García-Cepero, M. C., & McCoach, D. B. (2009). Educators' implicit theories of intelligence and beliefs about the identification of gifted students. *Universitas Psychologica, 8*(2), 295–310.

García-Ros, R., Talaya, I., & Pérez-González, F. (2012). The process of identifying gifted children in elementary education: Teachers' evaluations of creativity. *School Psychology International*, *33*, 661–672. http://doi.org/10.1177/0143034311421434

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Geake, J. G. (2009). Neuropsychological characteristics of academic and creative giftedness. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 261–273). Springer. http://doi.org/10.1007/978-1-4020-6162-2 11

Gear, G. H. (1976). Accuracy of teacher judgment in identifying intellectually gifted children: A review of the literature. *Gifted Child Quarterly*, *20*, 478–490. http://doi.org/10.1177/001698627602000416

Gear, G. H. (1978). Effects of training on teachers' accuracy in the identification of gifted children. *Gifted Child Quarterly, 22*, 90–97. http://doi.org/10.1177/001698627802 200121

Geiser, C., Mandelman, S. D., Tan, M., & Grigorenko, E. L. (2016). Multitrait-multimethod assessment of giftedness: An application of the Correlated Traits-Correlated (Methods-1) Model. *Structural Equation Modeling: A Multidisciplinary Journal, 23*, 76–90. http://doi.org/10.1080/10705511.2014.937792

Genser, B., Strasser, E.-M., & Garbe, U. (1981). Stimmen Mütter und Kinder in der Einschätzung ihres schulbezogenen Verhaltens überein? [Do mothers and children agree in their evaluation of their school-related behavior?]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 13*, 341–347.

George, D. M., & Mallery, P. C. (2014). IBM SPSS Statistics 21 step by step: A simple guide and reference (13th ed.). Boston, MA: Pearson.

Givvin, K. B., Stipek, D. J., Salmon, J. M., & MacGyvers, V. L. (2001). In the eyes of the beholder: Students' and teachers' judgments of students' motivation. *Teaching and Teacher Education, 17*, 321–331. http://doi.org/10.1016/S0742-051X(00)00060-3

Glueck, C. L., & Reschly, A. L. (2014). Examining congruence within school-family partnerships: Definition, importance, and current measurement approaches. *Psychology in the Schools, 51*, 296–315. http://doi.org/10.1002/pits.21745

Goodnough, K. (2000). Fostering liberal views of giftedness: A study of the beliefs of six undergraduate education students. Roeper Review, 23, 89–90. doi:10.1080/02783190009554073

Gottfried, A. W., Gottfried, A. E., & Guerin, D. W. (2006). The Fullerton Longitudinal Study: A long-term investigation of intellectual and motivational giftedness. *Journal for the Education of the Gifted, 29*, 430–450. http://doi.org/10.4219/jeg-2006-244

Gralewski, J., & Karwowski, M. (2013). Polite girls and creative boys? Students' gender moderates accuracy of teachers' ratings of creativity. *The Journal of Creative Behavior, 47*, 290–304. http://doi.org/10.1002/jocb.36

Grantham, T. C. (2012). Eminence-focused gifted education: Concerns about forward movement void of an equity vision. *Gifted Child Quarterly, 56*, 215–220. http://doi.org/10.1177/0016986212456074

Gröhlich, C., & Guill, K. (2009). How stable are reference group effects of secondary school track recommendations? *Journal for Educational Research Online, 1*, 154–171.

Gross, M. U. (1999). Small poppies: Highly gifted children in the early years. *Roeper Review, 21*, 207–214. http://doi.org/10.1080/02783199909553963

Cunningham, L. G., & Rinn, A. N. (2007). The role of gender and previous participation in a summer program on gifted adolescents' self-concepts over time. *Journal for the Education of the Gifted, 30*, 326–352. http://doi.org/10.1177/016235320703000303

Guskin, S. L., Peng, C.-Y. J., & Majd-Jabbari, M. (1988). Teachers' perceptions of giftedness. *Gifted Child Quarterly, 32*, 216–221. http://doi.org/10.1177/001698628803200106

Guskin, S. L., Peng, C.-Y. J., & Simon, M. (1992). Do teachers react to "multiple intelligences"? Effects of teachers' stereotypes on judgments and expectancies for students with diverse patterns of giftedness/talent. *Gifted Child Quarterly, 36*, 32–37. http://doi.org/10.1177/001698629203600108

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence, 45*, 34–45. http://doi.org/10.1016/j.intell.2013.04.001

Hansen, J. B., & Feldhusen, J. F. (1994). Comparison of trained and untrained teachers of gifted students. *Gifted Child Quarterly, 38*, 115–121. http://doi.org/10.1177/001698629403800304

Hany, E. A. (1993). How teachers identify gifted students: feature processing or concept based classification. *European Journal for High Ability, 4*, 196–211. http://doi.org/10.1080/0937445930040208

Hany, E. A. (1997). Modeling teachers' judgment of giftedness: a methodological inquiry of biased judgment. *High Ability Studies*, *8*, 159–178. http://doi.org/10.1080/1359813970080203

Harder, B., Trottler, S., Vialle, W., & Ziegler, A. (2015). Diagnosing resources for effective learning via teacher and parent checklists. *Psychological Test and Assessment Modeling, 57*, 201–221.

Harradine, C. C., Coleman, M. R. B., & Winn, D.-M. C. (2014). Recognizing academic potential in students of color findings of U~STARS~PLUS. *Gifted Child Quarterly, 58*, 24–34. http://doi.org/10.1177/0016986213506040

Hartwig, M., & Bond Jr, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643. http://doi.org/10.1037/a0023589

Harvey, K. E., Suizzo, M.-A., & Jackson, K. M. (2016). Predicting the grades of low-income-ethnic-minority students from teacher-student discrepancies in reported motivation. *The Journal of Experimental Education, 84*, 510–528. http://doi.org/10.1080/00220973.2015.1054332

Hauser-Cram, P., Sirin, S. R., & Stipek, D. (2003). When teachers' and parents' values differ: Teachers' ratings of academic competence in children from low-income families. *Journal of Educational Psychology, 95*, 813. http://doi.org/10.1037/0022-0663.95.4.813

Heinbokel, A. (2001). *Hochbegabung im Spiegel der Printmedien seit 1950. Vom Werdegang eines Bewusstseinswandels*. Bonn: BMBF Publik.

Heller, K. A. (2004). Identification of gifted and talented students. *Psychology Science*, *46*, 302-323.

Heller, K. A., & Perleth, C. (2007). *Münchner Hochbegabungstestbatterie für die Primarstufe (MHBT-P)* [Munich giftedness test battery for primary school students]. Goettingen: Hogrefe.

Heller, K. A., & Perleth, C. (2008). The Munich High Ability Test Battery (MHBT): A multidimensional, multimethod approach. *Psychology Science, 50*, 173.

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* (2nd ed., pp. 147–170). New York, NY: Cambridge University Press.

Heller, K. A., Reimann, R., & Senfter, A. (2005). *Hochbegabung im Grundschulalter: Erkennen und Fördern* [Giftedness in elementary school: Assessing and encouraging]. Münster: Lit.

Heller, K. A., & Schofield, N. J. (2008). Identifying and nurturing the gifted: An international perspective. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children. Psychoeducational theory, research, and best practices* (pp. 93–114). Springer.

Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* [Instructional quality and teacher professionality: Diagnosis, evaluation, and enhancement of instruction]. Seelze, Germany: Klett-Kallmeyer.

Helmke, A., & Schrader, F.-W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching and Teacher Education, 3*, 91–98. http://doi.org/10.1016/0742-051X(87)90010-2

Helmke, A., & Schrader, F.-W. (1989). Sind Mütter gute Diagnostiker ihrer Kinder? Analysen von Komponenten und Determinanten der Urteilsgenauigkeit [Are mothers good diagnosticians? Analysis of components and determinants of judgment accuracy]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 21*, 223-247.

Hernández-Torrano, D., Prieto, M. D., Ferrándiz, C., Bermejo, R., & Sáinz, M. (2013). Characteristics leading teachers to nominate secondary students as gifted in Spain. *Gifted Child Quarterly, 57*, 181–196. http://doi.org/10.1177/0016986213490197

Hickey, M. G., & Toth, L. (1990). The effects of labeling children gifted: A review of the literature. *Early Child Development and Care, 63*, 149–151. http://doi.org/10.1080/0300443900630118

Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology, 101*, 662. http://doi.org/10.1037/a0014306

Hoge, R. D. (1983). Psychometric properties of teacher-judgment measures of pupil aptitudes, classroom behaviors, and achievement levels. *The Journal of Special Education, 17*, 401–429. http://doi.org/10.1177/002246698301700404

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research, 59*, 297–313. http://doi.org/10.3102/00346543059003297

Hoge, R. D., & Cudmore, L. (1986). The use of teacher-judgment measures in the identification of gifted pupils. *Teaching and Teacher Education*, *2*, 181–196. http://doi.org/10.1016/0742-051X(86)90016-8

Hodge, K. A., & Kemp, C. R. (2006). Recognition of giftedness in the early years of school: Perspectives of teachers, parents, and children. *Journal for the Education of the Gifted, 30*, 164–204. http://doi.org/10.4219/jeg-2006-259

Hollingworth, L. S. (1942). *Children above 180 IQ Stanford-Binet: Origin and development*. New York, US: World Books.

Hong, E., Greene, M., & Hartzell, S. (2011). Cognitive and motivational characteristics of elementary teachers in general education classrooms and in gifted programs. *Gifted Child Quarterly, 55*, 250–264. http://doi.org/10.1177/0016986211418107

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

Hunsaker, S. L., Finley, V. S., & Frank, E. L. (1997). An analysis of teacher nominations and student performance in gifted programs. *Gifted Child Quarterly*, *41*, 19–24. http://doi.org/10.1177/001698629704100203

Hunt, J. M., & Paraskevopoulos, J. (1980). Children's psychological development as a function of the inaccuracy of their mothers' knowledge of their abilities. *The Journal of Genetic Psychology, 136*, 285–298. http://doi.org/10.1080/00221325.1980.10534123

Jacobs, J. C. (1971). Effectiveness of teacher and parent identification of gifted children as a function of school level. *Psychology in the Schools, 8*, 140–142. http://doi.org/10.1002/1520-6807(197104)8:2<140::AID-PITS2310080210>3.0.CO;2-K

Jacobs, J. C. (1972). Teacher attitude toward gifted children. *Gifted Child Quarterly, 16*, 23-26.

Jamieson-Noel, D., & Winne, P. H. (2003). Comparing self-reports to traces of studying behavior as representations of students' studying and achievement. *German Journal of Educational Psychology, 17*, 159-171. http://doi.org/10.1024/1010-0652.17.34.159

Jarosewich, T., Pfeiffer, S. I., & Morris, J. (2002). Identifying gifted students using teacher rating scales: A review of existing instruments. *Journal of Psychoeducational Assessment*, 20, 322–336.

Johnson, D. R., & Young, R. (2011). Toward best practices in analyzing datasets with missing data: Comparisons and recommendations. *Journal of Marriage and Family, 73*, 926–945. http://doi.org/10.1111/j.1741-3737.2011.00861.x

Jones, B. D., Bryant, L. H., Snyder, J. D., & Malone, D. (2012). Preservice and inservice teachers' implicit theories of intelligence. *Teacher Education Quarterly*, *39*, 87–101.

Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2012). Personality traits moderate the big-fish-little-pond effect of academic self-concept. *Learning and Individual Differences, 22*, 736–746. http://doi.org/10.1016/j.lindif.2012.07.020

Jonsson, A.-C., Beach, D., Korp, H., & Erlandson, P. (2012). Teachers' implicit theories of intelligence: Influences from different disciplines and scientific theories. *European Journal of Teacher Education, 35*, 387–400. http://doi.org/10.1080/02619768.2012.662636

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9*, 131–155. http://doi.org/10.1207/s15327957pspr0902_3

Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction, 28*, 73–84. http://doi.org/10.1016/j.learninstruc.2013.06.001

Karlson, K. B., Holm, A., & Breen, R. (2012). Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology*, *42*, 286–313. http://doi.org/10.1177/0081175012444861

Karnes, M. B., & Johnson, L. J. (1990). A plea: Serving young gifted children. *Early Child Development and Care, 63*, 131–138. http://doi.org/10.1080/0300443900630116

Kaufman, S. B., & Sternberg, R. J. (2008). Conceptions of giftedness. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children. Psychoeducational theory, research, and best practices* (pp. 71–91). New York, NY: Springer. http://doi.org/10.1007/978-0-387-74401-8_5

Keogh, B., & MacMillan, D. (1996). Exceptionality. In D. C. Berliner, & R. C. Calfee (Ed.), *Handbook of educational psychology* (pp. 311–330). NewYork: Simon Shuster New York.

Kim, M. (2016). A Meta-Analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly, 60*, 102 –116. http://doi.org/10.1177/0016986216630607

Kleber, E.W. (1992). Diagnostik in pädagogischen Handlungsfeldern. Einführung in Bewertung, Beurteilung, Diagnose und Evaluation [Diagnostic in educational areas of activity. An introduction to assessment, judgment, diagnostic, and evaluation]. Weinheim: Juventa Verlag.

Kolb, K. J., & Jussim, L. (1994). Teacher expectations and underachieving gifted children. *Roeper Review, 17*, 26–30. http://doi.org/10.1080/027831994095536 13

Kretschmann, J., Vock, M., & Lüdtke, O. (2014). Acceleration in elementary school: Using propensity score matching to estimate the effects on academic achievement. *Journal of Educational Psychology, 106*, 1080. http://doi.org/10.1037/a0036631

Kulik, J. A., & Kulik, C.-L. C. (1984). Effects of accelerated instruction on students. *Review of Educational Research, 54*, 409–425. http://doi.org/10.3102/003465430 54003409

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

Kultusministerkonferenz (2015). *Förderstrategie für leistungsstarke Schülerinnen und Schüler [Promotion strategies for high achieving students]. (Beschluss der Kultusministerkonferenz vom 11.06.2015)*. Retrieved from http://www.kmk.org/ fileadmin/pdf/350-KMK-TOP-011-Fu-Leistungsstarke_-_neu.pdf

Lee, L. (1999). Teachers' conceptions of gifted and talented young children. *High Ability Studies*, *10*, 183–196. http://doi.org/10.1080/1359813990100205

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research, 41*, 63–76. http://doi.org/10.1080/0013188990410106

Li, H., Lee, D., Pfeiffer, S. I., Kamata, A., Kumtepe, A. T., & Rosado, J. (2009). Measurement invariance of the Gifted Rating Scales—School Form across five cultural groups. *School Psychology Quarterly, 24*, 186–198. http://doi.org/10.1037/ a0017382

Li, H., Lee, D., Pfeiffer, S. I., & Petscher, Y. (2008). Parent ratings using the Chinese version of the Parent Gifted Rating Scales-School Form. Reliability and validity for Chinese students. *Educational and Psychological Measurement, 68*, 659–675.

Lipowsky, F., & Rzejak, D. (2012). Lehrerinnen und Lehrer als Lerner. Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen [Teachers as learners. When does the change of roles work? Characteristics and effects of effective teacher education programs]. *Schulpädagogik Heute, 3*, 1–17.

Lohman, D. (2005). *Identifying academically talented minority students*. Storrs: National Research Center on the Gifted and Talented.

Lohman, D. F., & Gambrell, J. L. (2012). Using nonverbal tests to help identify academically talented children. *Journal of Psychoeducational Assessment, 30*, 25–44. http://doi.org/10.1177/0734282911428194

Louis, B., & Lewis, M. (1992). Parental beliefs about giftedness in young children and their relation to actual ability level. *Gifted Child Quarterly, 36*, 27–31. http://doi.org/10.1177/001698629203600107

Lüdtke, O., Köller, O., Marsh, H. W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, *30*, 263-285. http://doi.org/10.1016/j.cedpsych.2004.10.002

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives, 2*, 99–106. http://doi.org/10.1111/j.1750-8606.2008.00048.x

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

Machts, N., Kaiser, J., Schmidt, F. T., & Möller, J. (2016). Accuracy of teachers' judgments of students' cognitive abilities: A meta-analysis. *Educational Research Review, 19*, 85–103. http://doi.org/10.1016/j.edurev.2016.06.003

Makel, M. C., Lee, S.-Y., Olszewki-Kubilius, P., & Putallaz, M. (2012). Changing the pond, not the fish: Following high-ability students across different educational environments. *Journal of Educational Psychology, 104*, 778. http://doi.org/10.1037/a0027558

Manaster, G. J., Chan, J. C., Watt, C., & Wiehe, J. (1994). Gifted adolescents' attitudes toward their giftedness: A partial replication. *Gifted Child Quarterly, 38*, 176–178. http://doi.org/10.1177/001698629403800404

Marsh, H. W. (1984). Self-concept: The application of a frame of reference model to explain paradoxical results. *Australian Journal of Education, 28*, 165-181. http://doi.org/10.1037/0003-066X.58.5.364

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*, 280–295. doi:0022-0663/87/S00.75

Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: The big fish strikes again. *American*

Educational Research Journal, 32, 285–319. http://doi.org/10.3102/00028312032002285

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology, 78*, 337–349. http://doi.org/10.1037//0022-3514.78.2.337

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the Big Five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471–491. http://doi.org/10.1037/a0019227

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110. http://doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Nagengast, B., & Morin, A. J. (2012). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*, 1194-1218. http://doi.org/10.1037/a0026913

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*, 213. http://doi.org/10.1037/0022-3514.47.1.213

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008). The Big-Fish-Little-Pond-Effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, *20*, 319-350. http://doi.org/10.1007/s10648-008-9075-6

Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions. A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development, 39*, 87–96. http://doi.org/10. 1177/0165025414552301

Mason, B. A., Gunersel, A. B., & Ney, E. A. (2014). Cultural and ethnic bias in teacher ratings of behavior: A criterion-focused review. *Psychology in the Schools, 51*, 1017–1030. http://doi.org/10.1002/pits.21800

Matthews, M. S., Ritchotte, J. A., & Jolly, J. L. (2014). What's wrong with giftedness? Parents' perceptions of the gifted label. *International Studies in Sociology of Education, 24*, 372–393. http://doi.org/10.1080/09620214.2014.990225

Mayer, R. E. (2005). The scientific study of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 437–447). New York: Cambridge University Press.

McBee, M. T. (2006). A descriptive analysis of referral sources for gifted identification screening by race and socioeconomic status. *Prufrock Journal, 17*, 103–111. http://doi.org/10.4219/jsge-2006-686

McBee, M. T. (2010). Examining the probability of identification for gifted programs for students in Georgia elementary schools: A multilevel path analysis study. *Gifted Child Quarterly, 54*, 283–297. http://doi.org/10.1177/0016986210377927

McBee, M. T., Peters, S. J., & Miller, E. M. (2016). The impact of the nomination stage on gifted program identification. A comprehensive psychometric analysis. *Gifted Child Quarterly, 60*, 258–278. http://doi.org/10.1177/0016986216656256

McCall, R. B. (1977). Childhood IQ's as predictors of adult educational and occupational status. *Science, 197*, 482–483. http://doi.org/10.1126/science.197.4302.482

McClain, M.-C., & Pfeiffer, S. (2012). Identification of gifted students in the United States today: A look at state definitions, policies, and practices. *Journal of Applied School Psychology, 28*, 59–88. http://doi.org/10.1080/15377903.2012.643757

McCoach, D. B., & Siegle, D. (2007). What predicts teachers' attitudes toward the gifted? *Gifted Child Quarterly, 51*, 246–254. http://doi.org/10.1177/0016986207302719

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10. http://doi.org/10.1016/j.intell.2008.08.004

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568–592. http://doi.org/10.1037/0021-9010.93.3.568

Megay-Nespoli, K. (2001). Beliefs and attitudes of novice teachers regarding instruction of academically talented learners. *Roeper Review, 23*, 178–182. http://doi.org/10.1080/02783190109554092

Mendaglio, S., & Tillier, W. (2006). Dabrowski's theory of positive disintegration and giftedness: Overexcitability research findings. *Journal for the Education of the Gifted, 30*, 68–87. http://doi.org/10.1177/016235320603000104

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543. http://doi.org/10.1007/BF02294825

Milek, A., Lüdtke, O., Trautwein, U., Maaz, K., & Stubbe, T. C. (2009). Wie konsistent sind Referenzgruppeneffekte bei der Vergabe von Schulformempfehlungen? Bundeslandspezifische Analysen mit Daten der IGLU-Studie [How consistent are reference group effects on teachers' recommendations at the transition to secondary education? State-specific analyses of PIRLS data]. In J. Baumert, K. Maaz, & U. Trautwein (Ed.), *Bildungsentscheidungen* (pp. 282–301). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften. http://doi.org/10.1007/978-3-531-92216-4_12

Miller, E. M. (2005). Studying the meaning of giftedness: Inspiration from the field of cognitive psychology. *Roeper Review*, *27*, 172–177. http://doi.org/10.1080/02783190509554312

Miller, E. M. (2009). The effect of training in gifted education on elementary classroom teachers' theory-based reasoning about the concept of giftedness. *Journal for the Education of the Gifted, 33*, 65–105. http://doi.org/10.1177/016235320903300104

Miller, S. A., & Davis, T. L. (1992). Beliefs about children: A comparative study of mothers, teachers, peers, and self. *Child Development, 63*, 1251–1265. http://doi.org/10.1111/j.1467-8624.1992.tb01693.x

Mönks, F. J., & Katzko, M. W. (2005). Giftedness and gifted education. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 187–200). New York, US: Cambridge University Press.

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

Moon, T. R., & Brighton, C. M. (2008). Primary teachers' conceptions of giftedness. *Journal for the Education of the Gifted*, *31*, 447–480.

Moulton, P., Moulton, M., Housewright, M., & Bailey, K. (1998). Gifted and talented: Exploring the positive and negative aspects of labeling. *Roeper Review, 21*, 153–154. http://doi.org/10.1080/02783199809553950

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998-2015). Mplus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316. http://doi.org/10.1234/12345678

National Association for Gifted Children. (2013). *2012-2013 State of the states in gifted education*. Washington, DC: Author.

Neber, H. (2004). Teacher identification of students for gifted programs: nominations to a summer school for highly-gifted students. *Psychology Science*, *46*, 348-362.

Nogueira, S. M. (2006). MORCEGOS: A Portuguese enrichment program of creativity pilot study with gifted students and students with learning difficulties. *Creativity Research Journal, 18*, 45–54. http://doi.org/10.1207/s15326934crj1801_6

O'Hara, L. A., & Sternberg, R. J. (2001). It doesn't hurt to ask: Effects of instructions to be creative, practical, or analytical on essay-writing performance and their interaction with students' thinking styles. *Creativity Research Journal, 13*, 197–210. http://doi.org/10.1207/S15326934CRJ1302_7

Olszewski-Kubilius, P., & Clarenbach, J. (2012). *Unlocking emergent talent: Supporting high achievement of low-income, high ability students*. Washington, DC: National Association for Gifted Children.

Olszewski-Kubilius, P., & Thomson, D. (2012). Gifted education programs and procedures. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 2, pp. 487–510). Hoboken, New Jersey: Wiley Online Library. http://doi.org/10.1002/9781118133880.hop207016

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*, 307–332. http://doi.org/10.3102/00346543062003307

Peet, S. H., Powell, D. R., & O'Donnel, B. K. (1997). Mother-teacher congruence in perceptions of the child's competence and school engagement: Links to academic achievement. *Journal of Applied Developmental Psychology, 18*, 373–393. http://doi.org/10.1016/S0193-3973(97)80006-8

Pegnato, C. V., & Birch, J. W. (1959). Locating gifted children in junior high schools: A comparison of methods. *Exceptional Children*, *25*, 300–304.

Persson, R. S. (1998). Paragons of virtue: Teachers' conceptual understanding of high ability in an egalitarian school system. *High Ability Studies, 9*, 181–196. http://doi.org/10.1080/1359813980090204

Peters, S. J., & Gentry, M. (2010). Multigroup construct validity evidence of the HOPE Scale: Instrumentation to identify low-income elementary students for gifted programs. *Gifted Child Quarterly, 54*, 298–313. http://doi.org/10.1177/0016986210378332

Peters, S. J., Matthews, M., McBee, M., & McCoach, D. B. (2014). *Beyond gifted education: Designing and implementing advanced academic programs*. Waco, Texas: Sourcebooks, Inc.

Petersen, J. (2013). Gender differences in identification of gifted youth and in gifted program participation: A meta-analysis. *Contemporary Educational Psychology, 38*, 342–348. http://doi.org/10.1016/j.cedpsych.2013.07.002

Peterson, J. S., & Margolin, L. (1997). Naming gifted children: An example of unintended "reproduction." *Journal for the Education of the Gifted, 21*, 82–100. http://doi.org/10.1177/016235329702100105

Petscher, Y., & Li, H. (2008). Measurement invariance of the Chinese Gifted Rating Scales. Teacher and parent forms. *Journal of Psychoeducational Assessment, 26*, 274–286. http://doi.org/10.1177/0734282907303873

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*, 751–783. http://doi.org/10.1037/0022-3514.90.5.751

Piopiunik, M., & Schlotter, M. (2012). Identifying the incidence of "grading on a curve": A within-student across-subject approach. *Ifo Working Paper 121*.

Pfeiffer, S. I. (2015). *Essentials of gifted assessment*. Hoboken, New Jersey: John Wiley.

Pfeiffer, S. I., & Jarosewich, T. (2003). *Gifted rating scales*. San Antonio, TX: The Psychological Corporation.

Pfeiffer, S. I., Petscher, Y., & Kumtepe, A. (2008). The Gifted Rating Scales-School Form: A validation study based on age, gender, and race. *Roeper Review, 30*, 140–146. http://doi.org/10.1080/02783190801955418

Plucker, J. A., & Barab, S. A. (2005). The importance of contexts in theories of giftedness. Learning to embrace the messy joys of subjectivity. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 201–216). New York, US: Cambridge University Press.

Plucker, J. A., & Callahan, C. M. (2012). Introduction to the Special Issue. *Gifted Child Quarterly, 56*, 175. http://doi.org/10.1177/0016986212456078

Plucker, J. A., & Callahan, C. M. (2014). Research on giftedness and gifted education status of the field and considerations for the future. *Exceptional Children, 80*, 390–406. http://doi.org/10.1177/0014402914527244

Preckel, F., & Brüll, M. (2010). The benefit of being a big fish in a big pond: Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences, 20*, 522–531. http://doi.org/10.1016/j.lindif.2009.12.007

Preckel, F., & Vock, M. (2013). *Hochbegabung: Ein Lehrbuch zu Grundlagen, Diagnostik und Fördermöglichkeiten* [Giftedness: A textbook on basics, diagnostic, and promotion possibilities]. Hogrefe Verlag.

Rambo, K. E., & McCoach, D. B. (2012). Teacher attitudes toward subject-specific acceleration: Instrument development and validation. *Journal for the Education of the Gifted, 35*, 129–152.

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok—Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*, 731–737. http://doi.org/10.1016/j.jesp.2011.12.012

Rauer, W., & Schuck, K. D. (2003). *FEESS 3 – 4. Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen* [FEESS 3 – 4. Questionnaire for Assessing Emotional and Social School Experiences of 3rd and 4th Grade Primary School Children]. Göttingen: Hogrefe.

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*, 233–235. http://doi.org/10.1177/0016986209346824

Rennen-Allhoff, B. (1991). Wie verläßlich sind Elternangaben? [How reliable are parents' reports?]. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 40*, 333–338.

Renzulli, J. S. (2005b). *Equity, excellence, and economy in a system for identifying students in gifted education programs: A guidebook* (RM05208). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.

Renzulli, J. S. (2005a). The three-ring conception of giftedness. A developmental model for promoting creative productivity. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 246–279). New York, US: Cambridge University Press.

Renzulli, J. S., & Delcourt, M. A. B. (1986). The legacy and logic of research on the identification of gifted persons. *Gifted Child Quarterly, 30*, 20–23. http://doi.org/10.1177/001698628603000104

Robinson, N. M. (2005). In defense of a psychometric approach to the definition of academic giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 280-294). New York: Cambridge University Press.

Robinson, N. M., Zigler, E., & Gallagher, J. J. (2000). Two tails of the normal curve: Similarities and differences in the study of mental retardation and giftedness. *American Psychologist, 55*, 1413.

Rogers, K. B. (1986). Do the gifted think and learn differently? A review of recent research and its implications for instruction. *Journal for the Education of the Gifted, 10*, 17–39.

Rohrer, J. C. (1995). Primary teacher conceptions of giftedness: Image, evidence, and nonevidence. *Journal for the Education of the Gifted, 18*, 269–283. http://doi.org/10.1177/016235329501800304

Ross, P. O. (1993). *National excellence: A case for developing America's talent*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Rost, D. H. (2008). Multiple Intelligenzen, multiple Irritationen [multiple intelligences, multiple irritations]. *Zeitschrift Für Pädagogische Psychologie, 22*, 97–112. http://doi.org/10.1024/1010-0652.22.2.97

Rost, D. H. (2013). *Handbuch Intelligenz [Handbook of intelligence]*. Weinheim, Germany: Beltz.

Rost, D. H., & Hanses, P. (1997). Wer nichts leistet, ist nicht begabt? Zur Identifikation hochbegabter Underachiever durch Lehrkräfte [Not achieving not gifted? About the identification of gifted underachievers by teacher ratings]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 29*, 167–177.

Rost, D. H., & Schilling, S. R. (2006). Hochbegabung [Giftedness]. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (3rd ed., pp. 233–243). Weinheim: Beltz.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137. http://doi.org/10.1016/j.intell.2015.09.002

Rothenbusch, S., Zettler, I., Lösch, T., & Voss, T. (2016). *Elementary school teachers', enrichment program teachers', and student teachers' beliefs about the nature of giftedness.* Manuscript submitted for publication.

Rothenbusch, S., Zettler, I., Voss, T., Lösch, T., & Trautwein U. (2016). Exploring reference group effects on teachers' nominations of gifted students. *Journal of Educational Psychology, 108, 883–897.* http://doi.org/10.1037/edu0000085

Rubenzer, R. L., & Twaite, J. A. (1979). Attitudes of 1,200 educators toward the education of the gifted and talented: Implications for teacher preparation. *Journal for the Education of the Gifted, 2*, 202–213.

Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology, 80*, 121–135. http://doi.org/10.1348/000709909X466334

Runco, M. A. (1989). Parents' and teachers' ratings of the creativity of children. *Journal of Social Behavior and Personality, 4*, 73–83.

Runco, M. A. (1993). Divergent thinking, creativity, and giftedness. *Gifted Child Quarterly, 37*, 16–22. http://doi.org/10.1177/001698629303700103

Şahin, F., & Çetinkaya, Ç. (2015). An investigation of the effectiveness and efficiency of classroom teachers in the identification of gifted students. *Turkish Journal of Giftedness and Education, 5*, 133–146.

Şahin, N., & Düzen, E. (1994). The gifted child stereotype among university students and elementary school teachers. In K. A. Heller & E. A. Hany and (Eds.), *Competence and responsibility: Proceedings of the 3rd ECHA conference* (Vol. 2, pp. 367-376). Göttingen: Hogrefe & Huber Publishers.

Schack, G. D., & Starko, A. J. (1990). Identification of gifted students: An analysis of criteria preferred by preservice teachers, classroom teachers, and teachers of the gifted. *Journal for the Education of the Gifted*, *13*, 346–363. http://doi.org/10.1177/016235329001300405

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*, 1–10. http://doi.org/10.1037/a0018082

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*, 90. http://doi.org/10.1037/a0015108

Schneider, W., Niklas, F., & Schmiedeler, S. (2014). Intellectual development from early childhood to early adulthood: The impact of early IQ differences on stability and change over time. *Learning and Individual Differences, 32*, 156–162. http://doi.org/10.1016/j.lindif.2014.02.001

Schneider, W., Preckel, F., & Stumpf, E. (2014). *Hochbegabtenförderung in der Sekundarstufe*. Karg-Stiftung.

Schofield, N. J. & Hotulainen, R. (2004). Does all cream rise? The plight of unsupported gifted children. *Psychology Science, 46*, 379–386.

Schrader, F.-W. (2010). Diagnostische Kompetenz von Eltern und Lehrern [Diagnostic competence of parents and teachers]. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (4th ed., pp. 102-108). Weinheim: Beltz.

Schroth, S. T., & Helfer, J. A. (2009). Practitioners' conceptions of academic talent and giftedness: Essential factors in deciding classroom and school composition. *Journal of Advanced Academics*, *20*, 384–403. http://doi.org/10.1177/1932202X0902000302

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*, 455–498. http://doi.org/10.2307/1170362

Shavinina, L. V. (2009). *International handbook on giftedness*. Springer. http://doi.org/10.1007/978-1-4020-6162-2

Seaton, M., Marsh, H. W., & Craven, R. G. (2010). Big-Fish-Little-Pond Effect: Generalizability and moderation—two sides of the same coin. *American Educational Research Journal, 47*, 390–433. http://doi.org/10.3102/0002831209350493

Siegle, D. (2001). Overcoming bias in gifted and talented referrals. *Gifted Education Communicator, 32*, 22-25.

Siegle, D., Moore, M., Mann, R. L., & Wilson, H. E. (2010). Factors that influence in-service and preservice teachers' nominations of students for gifted and talented programs. *Journal for the Education of the Gifted, 33*, 337–360.

Siegle, D., & Powell, T. (2004). Exploring teacher biases when nominating students for gifted programs. *Gifted Child Quarterly, 48*, 21–29. http://doi.org/10.1177/001698620404800103

Siegle, D., & Reis, S. M. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly, 42*, 39–47. doi:10.1177/001698629804200105

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., … Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 68. http://doi.org/10.1037/1931-3896.2.2.68

Simonton, D. K. (2005). Genetics of giftedness. The implications of an emergenic–epigenetic model. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 312–326). New York, US: Cambridge University Press.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection. *Educational and Psychological Measurement, 69*, 493–525. http://doi.org/10.1177/0013164408323233

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis - an introduction to basic and advanced multilevel modeling* (2 ed.). London: Sage.

Sommer, U., Fink, A., & Neubauer, A. C. (2008). Detection of high ability children by teachers and parents: Psychometric quality of new rating checklists for the assessment of intellectual, creative and social ability. *Psychology Science Quarterly, 50*, 189–205.

Southern, W. T., & Jones, E. D. (2004). Types of acceleration: Dimensions and issues. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived: How schools hold back America's brightest students* (Vol. 2, pp. 5–12). Iowa City, IA.

Speirs Neumeister, K. L., Adams, C. M., Pierce, R. L., Cassady, J. C., & Dixon, F. A. (2007). Fourth-grade teachers' perceptions of giftedness: Implications for identifying and serving diverse gifted students. *Journal for the Education of the Gifted, 30*, 479–499. http://doi.org/10.4219/jeg-2007-503

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz [Accuracy of teacher judgments on student characteristics and the construct of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie, 19*, 85–95. http://doi.org/10.1024/1010-0652.19.1.85

Spinath, B., & Spinath, F. M. (2005). Development of self-perceived ability in elementary school: The role of parents' perceptions, teacher evaluations, and intelligence. *Cognitive Development, 20*, 190–204. http://doi.org/10.1016/j.cogdev.2005.01.001

Steenbergen-Hu, S., & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly, 55*, 39–53. http://doi.org/10.1177/0016986210383155

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49, 607–627.

Sternberg, R. J. (1999). A triarchic approach to the understanding and assessment of intelligence in multicultural populations. *Journal of School Psychology, 37*, 145–159. http://doi.org/10.1016/S0022-4405(98)00029-6

Sternberg, R. J. (2005). The WICS model of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 327–243), New York: Cambridge University Press.

Sternberg, R. J., & Davidson, J. E. (2005), *Conceptions of giftedness* (2nd ed.). New York: Cambridge University Press.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2011). *Explorations in giftedness*. New York, NY: Cambridge University Press.

Sternberg, R. J., & Zhang, L. (1995). What do we mean by giftedness? A pentagonal implicit theory. *Gifted Child Quarterly*, *39*, 88–94. http://doi.org/10.1177/001698629503900205

Stoeger, H. (2009). The history of giftedness research. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 17–38). Springer Netherlands. http://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, *12*, 3–54. http://doi.org/10.1177/1529100611418056

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2012). A proposed direction forward for gifted education based on psychological science. *Gifted Child Quarterly*, *56*, 176–188. http://doi.org/10.1177/0016986212456079

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*, 743–762. http://doi.org/10.1037/a0027627

Südkamp, A., & Möller, J. (2009). Referenzgruppeneffekte im Simulierten Klassenraum [The simulated classroom: An experimental study on diagnostic competence]. *German Journal of Educational Psychology*, *23*, 161–174. http://doi.org/10.1024/1010-0652.23.34.161

Terman, L. M. (1925). *Genetic studies of genius: Vol. 1. Mental and physical traits of a thousand gifted children*. California: Stanford University Press.

Terman, L. M. (1954). The discovery and encouragement of exceptional talent. *American Psychologist*, *9*, 221. http://doi.org/10.1037/h0060516

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*, 25–29. http://doi.org/10.1037/h0071663

Thurstone, L. M. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Tiedemann, J., & Billmann-Mahecha, E. (2007). Zum Einfluss von Migration und Schulklassenzugehörigkeit auf die Übergangsempfehlung für die Sekundarstufe I [The influence of ethnic criteria and frame of reference effects on teachers' recommendations in regard to transition from primary to secondary education]. *Zeitschrift Für Erziehungswissenschaft*, *10*, 108–120. http://doi.org/10.1007/s11618-007-0009-8

Timmermans, A. C., de Boer, H., & van der Werf, M. P. (2016). An investigation of the relationship between teachers' expectations and teachers' perceptions of student attributes. *Social Psychology of Education, 19*, 1–24. http://doi.org/10.1007/s11218-015-9326-6

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology, 34*, 414-424.

Trautwein, U., & Baeriswyl, F. (2007). Wenn leistungsstarke Klassenkameraden ein Nachteil sind: Referenzgruppeneffekte bei Übergangsentscheidungen [When high-achieving classmates put students at a disadvantage: Reference group effects at the transition to secondary schooling]. *German Journal of Educational Psychology, 21*, 119-133. http://doi.org/10.1024/ 1010-0652.21.2.119

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*, 788–806. http://doi.org/10.1037/0022-0663.98.4.788

Urhahne, D. (2011). Teachers' judgments of elementary students' ability, creativity and task commitment. *Talent Development and Excellence, 3*, 229–237.

VanTassel-Baska, J. (2005). Domain-specific giftedness. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 358–376). New York, US: Cambridge University Press.

VanTassel-Baska, J., Feng, A. X., & Evans, B. L. (2007). Patterns of identification and performance among gifted students identified through performance tasks a three-year analysis. *Gifted Child Quarterly, 51*, 218–231. http://doi.org/10.1177/001698 6207302717

Van Yperen, N. W. (2009). Why some make it and others do not: Identifying psychological factors that predict career success in professional adult soccer. *Sport Psychologist, 23*, 317–329.

Vaughn, V. L., Feldhusen, J. F., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, *35*, 92–98. http://doi.org/10.1177/001698629103500208

Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2013). Mathematics beliefs. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 249–271). New York, US: Springer. http://doi.org/10.1007/978-1-4614-5149-5_12

Wagner, W., Helmke, A., & Schrader, F.-W. (2009). Die Rekonstruktion der Übergangsempfehlung für die Sekundarstufe I und der Wahl des Bildungsgangs auf der Basis des Migrationsstatus, der sozialen Herkunft, der Schulleistung und schulklassenspezifischer Merkmale [Reconstruction of recommendations and choices at the transition to secondary education: Ethnic criteria, social background, achievement, and class composition]. In J. Baumert, K. Maaz, & U. Trautwein

(Eds.), *Bildungsentscheidungen* (pp. 183–204). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften. http://doi.org/10.100 7/978-3-531-92216-4_8

Wahl, D., Huber, G., & Weinert, E. (2006). *Psychologie für die Schulpraxis* [Psychology for the school practice]. Bielefeld: Sozio Publishing.

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology, 102*, 860–871. http://doi.org/10.1037/a0019454

Watt, H. M. G., & Richardson, P. W. (2014). A motivational analysis of teacher beliefs. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 191-211). New York: Routledge.

Webb, J. T. (1994). *Nurturing social emotional development of gifted children*. ERIC Clearinghouse Reston, VA.

Westberg, K. L. (2012). Using teacher rating scales in the identification of students for gifted services. In S. L. Hunsaker (Ed.), *Identification: The theory and practice of identifying students for gifted and talented education services* (pp. 283-335). Mansfield Center, CT: Creative Learning Press.

Wild, K. P. (1993). Hochbegabtendiagnostik durch Lehrer [Giftedness diagnostic by teachers]. In D. H. Rost (Ed.), Lebensumweltanalyse hochbegabter Kinder: Das Marburger Hochbegabtenprojekt (pp. 236–261). Göttingen: Hogrefe.

Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8–10)* [Berlin Test of Fluid and Crystallized Intelligence for Grades 8–10]. Göttingen: Hogrefe.

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*, 242–244. http://doi.org/10.1177/0016 986209346828

Worrell, F. C., & Erwin, J. O. (2011). Best practices in identifying students for gifted and talented education programs. *Journal of Applied School Psychology*, *27*, 319–340. http://doi.org/10.1080/15377903.2011.615817

Worrell, F. C., Olszewski-Kubilius, P., & Subotnik, R. F. (2012). Important issues, some rhetoric, and a few straw men. A response to comments on “Rethinking Giftedness and Gifted Education." *Gifted Child Quarterly, 56*, 224–231. http://doi.org/10.1177/0016986212456080

Zeidner, M., & Schleyer, E. J. (1999). The big-fish-little-pond effect for academic self-concept, test anxiety, and school grades in gifted children. *Contemporary Educational Psychology, 24*, 305–329. http://doi.org/10.1006/ceps.1998.0985

Zettler, I., Thoemmes, F., Hasselhorn, M., & Trautwein, U. (2014). Aktuelle Lösungsansätze für methodische Herausforderungen in der Talent- und Begabungsforschung [Current solution approaches for methodological challenges in research on giftedness]. In M. Stamm, & A. Pfaller (Eds.), *Handbuch Talententwicklung: Theorien, Methoden und Praxis in Psychologie und Pädagogik* (pp. 475-484). Bern: Verlag Hans Huber Hogrefe AG.

Zhang, L., & Hui, S. K. (2002). From pentagon to triangle: A cross-cultural investigation of an implicit theory of giftedness. *Roeper Review*, *25*, 78–82. http://doi.org/10.1080/02783190309554203

Zhang, L., & Sternberg, R. J. (1998). The pentagonal implicit theory of giftedness revisited: A cross-validation in Hong Kong. *Roeper Review*, *21*, 149–153. http://doi.org/10.1080/02783199809553949

Ziegler, A. (2005). The Actiotope Model of Giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 411–436). New York: Cambridge University Press.

Ziegler, A., Kuhn, C., & Heller, K. A. (1998). Implizite Theorien von gymnasialen Mathematik- und Physiklehrkräften zu geschlechtsspezifischer Begabung und Motivation [Implicit theories of mathematics and physics teachers on gender-specific giftedness and motivation]. *Psychologische Beiträge, 40*, 271-287.

Ziegler, A., & Phillipson, S. N. (2012). Towards a systemic theory of gifted education. *High Ability Studies, 23*, 3–30. http://doi.org/10.1080/13598139.2012.679085

Ziegler, A., & Raul, T. (2000). Myth and reality: A review of empirical studies on giftedness. *High Ability Studies, 11*, 113–136. http://doi.org/10.1080/13598130020001188

Ziegler, A., Stoeger, H., Harder, B., & Balestrini, D. P. (2013). Gifted education in German-speaking Europe. *Journal for the Education of the Gifted, 36*, 384–411. http://doi.org/10.1177/0162353213492247

Ziegler, A., Stoeger, H., & Vialle, W. (2012). Giftedness and gifted education. The need for a paradigm change. *Gifted Child Quarterly, 56*, 194–197. http://doi.org/10.1177/0016986212456070