



Motivational and personality variables distinguish academic underachievers from high achievers, low achievers, and overachievers

Alexander Dings¹ · Frank M. Spinath¹

Received: 25 February 2021 / Accepted: 10 August 2021 / Published online: 4 September 2021
© The Author(s) 2021

Abstract

Previous studies have associated several variables concerning motivation and other domains with underachievement, i. e. a student's academic achievement falling short of what their cognitive abilities, as the best predictor of academic performance, would indicate. The present study extends these findings using a more rigorous approach in defining underachievers and suitable control groups. Using discriminant analysis, underachievers identified in a German twin family study were compared not only to achievers with comparable IQ scores, but also with students of lower aptitude, but comparably low grades, as well as overachieving students. Results confirm previous findings that compared to successful students, underachievers report lower levels of motivation and parental support; beyond this comparison, underachievers also differed from other low achievers, mostly in terms of their personality. In total, 40% of the variance between the groups were explained. Additionally, the data shed doubt on the common assumption that underachievers are an unusually heterogenous group of students.

Keywords Underachievement · Academic achievement · Personality · Motivation · Discriminant analysis

1 Introduction

1.1 General background

General intelligence has long been known as a major predictor of academic achievement (Sternberg, Grigorenko & Bundy, 2001), with a recent meta-analysis reporting a median correlation of 0.54 (Roth et al., 2015). Yet even this

✉ Alexander Dings
alexander.dings@uni-saarland.de

¹ Department of Psychology, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany

substantial predictive validity, which was corrected for sampling error, error of measurement and range restriction, implies that about three quarters of the variance in the outcome are still left unexplained. A substantial body of research has therefore attempted to uncover other predictors of educational attainment.

There are several approaches to this question. Schneider and Preckel (2017) reported effect sizes for more than 100 predictors in a synthesis of meta-analyses, integrating about 3300 single findings. While the resulting „ranking“ can hardly be taken literally as the interrelatedness between predictors is not taken into account, it illustrates clearly that influences on students' grades range across a variety of domains, such as personal traits, school variables, parental behavior, peers, etc.

Other studies have therefore attempted to investigate which of the variables in this heterogeneous mix really matter. A straightforward approach for identifying predictors that matter beyond intelligence is to simply control for candidate variables statistically. Different constructs have been found to show some incremental validity, including traits like conscientiousness (Rimfeld et al., 2016), self-efficacy (Stankov & Lee, 2014) and self-control (Duckworth et al., 2012).

Underachievement research asks a similar, but not identical question to address some of the 75% of unexplained variance left by IQ. It focuses on those individuals for whom IQ turns out to be off the mark in predicting achievement. With a less than perfect correlation between IQ and achievement, it is self-evident that there will necessarily be a certain proportion of individual cases in which the actual outcome will differ substantially from the predicted one. Some students will receive very good grades despite unassuming scores on general intelligence tests; some students will show above-average cognitive abilities when tested, but not do as well in school. These latter students have been termed "underachievers" in educational psychology, and in particular in giftedness research (for an overview, see Reis & McCoach, 2000).

By focusing on the subgroup of students in which the best overall predictor of academic success is notably off the mark, this approach attempts to further qualify the more general findings cited above since validity coefficients obtained in a larger population do not necessarily translate into every student's individual achievement history. The question that underachievement research asks could thus be phrased as: In the unusual case that it is not (low) IQ that explains a case of low achievement, what can explain it? It is not a given that factors known to be incrementally valid in general are the same that are at work in these "extreme" cases.

Previous studies have yielded a diverse range of findings on underachievement (see below). Its reverse case of overachievement has also been studied, albeit to a much lesser extent (Sparfeldt, Buch & Rost, 2010)—it might simply be a less pressing issue; also note that if a good IQ test result is used as a benchmark which the academic achievement of underachieving students is said to fall short of, you might just as well ask the reverse question for apparent "overachievers": why did their IQ test performance fall short of what they achieve in school? Considering that grades are an aggregated measure of performance throughout a school year, a comparatively

low result on a single IQ test might often be a fluke, resulting from random factors on that single test occasion.¹

However, there are two issues in the conceptual and statistical approaches widely used in these studies that, in our view, entail validity problems: Underachievers are rarely compared to other low achievers, making it unclear whether their characteristics are actually specific to achieving lower *than expected*; and the most widely used approach to operationalize underachievement leads to questionable selections. This study is aimed at testing and extending previous findings on underachievement while addressing these issues.

1.2 Operationalizing underachievement

In attempting to operationalizing underachievement, researchers have taken several different approaches with different implications (Lau & Chan, 2001). Numerous authors stress the fact that there is no consensus on how to conceptualize and measure underachievement (e. g., Hoover-Schultz, 2005; Plewis, 1991; Preckel et al., 2006). The definition of underachievement provided by Reis and McCoach (2000), according to which underachievers are “students who exhibit a severe discrepancy between expected achievement (as measured by standardized achievement test scores or cognitive or intellectual ability assessments) and actual achievement (as measured by class grades and teacher evaluations)” (p. 157), can be seen as common ground, but it leaves much room for different approaches in defining a “severe discrepancy”. We will therefore describe the approach that seems most rigorous and valid to us.

Much of the previous literature adopted either the *difference method* or the *regression method* (s. Lau & Chan, 2001) to compare standardized measures of cognitive abilities and achievement scores—either by subtracting them from one another or, in order to avoid regression to the mean (s. Cone & Wilson, 1981), by calculating the residuals in a regression of the achievement measure on the aptitude measure. The resulting difference or residuals are then interpreted as reflecting underachievement, and a cut-off is defined to select the most substantial cases. As a more recent, refined implementation of this rationale, Rasch models have been used, e. g. by Phillipson (2008). A different method, the *absolute split approach*, identifies underachievers categorically by defining cut-offs for low and high IQ as well as low and high grades, with the combination of high IQ and low grades representing underachievement. This approach might seem unfavorable at first glance because it dichotomizes the data and thereby discards some of the variance, but there are several advantages to it.

First of all, discarding some of the variance is not necessarily unwelcome. As alluded to in the introduction, the premise of underachievement research is to find out which (possibly unique) processes can explain those cases in which IQ as the

¹ As far as overt difficulties in administering the test are concerned, IQ results are excluded from analyses in TwinLife for those cases where interviewers reported any issues, thereby reducing the likelihood of IQ results not reflecting actual abilities.

Table 1 Framework of the absolute split approach for defining underachievement and corresponding comparison groups

	High IQ	Average IQ	Low IQ
High grades	(Expected) high achievers	<i>unselected</i>	Overachievers
Average grades	<i>unselected</i>	<i>unselected</i>	<i>unselected</i>
Low grades	Underachievers	<i>unselected</i>	(Expected) low achievers

most important predictor and academic achievement as the criterion do not align—in other words, extreme cases. The large middle ground of students who technically under- and overachieve to a small degree is not only of little interest in practice and hardly interpretable; these cases are also very prone to emerging simply due to random error variance, as the residuals in the regression method result from combining two variables with less than perfect reliability, making them rather noisy in a statistical sense (Zieger et al., 2012).

Second, the absolute split approach reflects a clear-cut definition of underachievement. As Hanses and Rost (1998) point out, this is questionable when using the regression approach, as it subsumes very different cases of partially unclear validity under the label of underachievement: it will „diagnose“ underachievement in many cases where the actual grades are just fine—but the student would have been expected to achieve even better results. While it can certainly be argued that this outcome might be a disappointment for the gifted student, we follow Sparfeldt, Schilling and Rost (1996) in reasoning that it would be dubious to use one and the same label for these students as for those whose unexpectedly low grades are actually critically low, and that an upper bound on grades is therefore advisable. Similarly, a lower bound for IQ is also advisable because it would be equally debatable to include students in the category of underachievement who were not expected to do all that well in the first place (such a lower bound is in place in most studies as they focused on gifted students). The absolute split method thereby focuses on those cases where there can be little disagreement about their validity (barring the possibility that they might have resulted from error variance): instances of underachievement are cases of academic failure (or being on the brink of failure) by students who score clearly above average in cognitive abilities. Note that such dramatic cases will be rarer in populations where the ability criterion and the achievement outcome are highly correlated; the regression method with 1 SD as the critical difference will, implausibly, always select about 15% of students as underachievers, regardless of whether IQ and grades correlate at 0.2 or at 0.7 (Plewis, 1991).²

² To illustrate, imagine a fictitious population where a person's cognitive abilities and their academic achievement are highly correlated because all factors that can hinder a student to make full use of their abilities have been eliminated. As long as the correlation is not exactly 1, the regression method will still identify 15% of students as underachievers and thus suggest that the issue persists; in fact, such a misunderstanding can sometimes be found in the literature (e. g., Figg, Rogers, McCormick and Low, 2012) when it is suggested that underachievement „rates“ have stayed at an alarmingly high level, when this is in fact simply inevitable under the most prevalent definition, as pointed out by McCall (1994).

Using the absolute split method offers the additional advantage of yielding convenient comparison groups (see Table 1): While the combination of high IQ and low grades represents underachievement, there are also groups combining high IQ and high grades (high achievers), low IQ and low grades (low achievers) and low IQ and high grades (overachievers).

This allows us to contrast underachievers with similar students differing either in grades or in IQ. Previous research mostly undertook only the first of these comparisons (i. e., the left column in Table 1) by focusing on gifted students and then contrasting underachievers with high achievers within this group, defined by the regression approach. This, of course, ensures that characteristics found in underachievers are not confounded with IQ (White et al., 2018). However, as McCall and Lau (2000) pointed out, this leaves a similar issue unresolved—„virtually all of the research (...) confounds these correlates with poor grades “ (p. 787). In consequence, it remains unclear whether differences between achieving and underachieving smart students are actually attributable to *underachievement* rather than to low achievement in general. Low academic self-concept, for example, might generally be prevalent among students who receive bad grades – they might, almost trivially, simply report this lack of success when asked about their academic self-concept.

Comparing underachievers simultaneously in both directions – contrasting them to other smart students as well as other unsuccessful students – therefore seems necessary to identify which variables typically associated with underachievement, such as low academic self-concept, are actually specific to it. Using the label “underachievement” in research and in practice implies that these students are in some way different from low achievers in general, but this assumption has hardly been tested empirically. To our knowledge, only two studies have addressed the confounding of underachievement research results by grades, even though McCoach and Siegle (2003) mentioned it as a suggestion for future research. McCall and Lau (2000) did so by comparing underachievers to non-underachievers while using grades as a covariate, but underachievement was defined by the regression approach, incurring the aforementioned issue of subsuming very different cases under one label. Hanses and Rost (1998) included several groups in their analysis and were thereby partly able to differentiate the results along grades and IQ, but their comparison groups included average achievers rather than low achievers. Other studies included non-gifted students as a control group, but did not differentiate them by grades (e. g., Vlahovic-Stetic, Vidosic and Arambasic, 1999; Dixon et al., 2006).

1.3 Empirical background

The research detailing the differences between underachievers and other students—usually gifted achievers—is diverse and sometimes contradictory (Peterson & Colangelo, 1996; Reis & McCoach, 2000; White et al., 2018) to such an extent that Figg, Rogers, McCornick, and Low (2012) conclude that there is „very little, if any, consensus on any answers“ (p. 54).

Some longitudinal research gives credence to the belief that the basic premise of underachievement research is valid: it is not mainly error variance that studies

are trying to explain; underachievement is a meaningful phenomenon with a possibly lasting impact. In other words, temporary underachievement does not simply “regress to the mean” in most cases, as one would expect if the discrepancies between grades and cognitive abilities were entirely the result of measurement error. For example, Colangelo, Kerr, Christensen and Maxey (1993) reported lower college admission test scores for high-school underachievers (although to a lesser extent than their underachievement would have suggested, indicating that some regression to the mean did occur). Studies by Johnson et al. (2006) and Peterson (2000) also indicate that at least some cases of underachievement persist longitudinally. On the other hand, all of these results also demonstrate that underachievement is clearly not an irreversible “condition”, as many students get back on track later on (see also Uno, Mortimer, Kim & Vuolo, 2010). Support from family and peers, as well as persistence in setting ambitious goals despite the difficulties, appear to be crucial factors for this to happen (Peterson, 2000; Hwang et al., 2014; McCall, 1994), making it all the more important to recognize underachievement as such and not mistake it for an unalterable lack of ability. This is not a trivial task: Rost and Hanses (1997) found that teachers in their study could not identify underachieving students in their classes very well; in a more recent study, teachers had only a moderately accurate impression of their students’ attitudes towards school—in particular, they often underestimated their students’ goal valuation (Siegle et al., 2020).

There is no clear overall picture on what determines underachievement, partly due to a relative paucity of comparable studies (White et al., 2018). Different procedural explanations for the development of underachievement have been suggested (e.g., Snyder & Linnenbrink-Garcia, 2013), but the empirical basis for them is scarce. The same is true for various “typologies” of underachievement suggested by practitioners (see Reis & McCoach, 2000). Nothing demonstrates this need for clarification better than the fact that interventions aiming to reduce underachievement have so far met with limited success overall, according to a recent meta-analysis (Steenbergen-Hu et al., 2020).

In the current study, we focus on two particular domains: motivation and personality.

1.3.1 Motivation

Motivational variables are at the core of one of the few theoretical models that attempt to provide a general explanation for at least some cases of underachievement. The Achievement Orientation Model (or AOM; Siegle et al., 2017), which is based on existing theories of motivation, hypothesizes that students’ motivational attitudes toward achievement serve as a necessary link between their abilities and actual success. In particular, (gifted) students need to believe that trying to be successful is a worthwhile goal (goal valuation); they need to believe in their ability to do so (self-efficacy); and they need to feel supported in their efforts by their environment, such as parents or teachers (environmental perceptions). The AOM presumes that these components cannot compensate for each other: each of them needs to be fulfilled at least to a minimal degree for the student to be successful. If this motivational basis is present, a behavioral component summarized as self-regulation is

further necessary to successfully put the motivation to succeed into practice; if the student does not know how to learn effectively, for example, their abilities and motivation will not translate into success.

It has repeatedly been found that the constructs referred to in the AOM are implicated in underachievement in the sense that compared (only) to high achievers, underachievers report lower academic self-concept, goal orientation and support from their parents (Balduf, 2009; Landis & Reschly, 2013; McCoach & Siegle, 2003; Mofield & Parker Peters, 2019; Vlahovic-Stetic et al., 1999; White et al., 2018). The AOM's other notion—that motivation must be complemented by skills and strategies that enable make it possible to learn efficiently—is also supported by empirical evidence. Hyperactivity and other kinds of dysfunctional behavior is an antecedent to underachievement in kids (as well as a consequence of it; Hinshaw, 1992). Castejon, Gilar, Veas & Miñano (2016) found lower use of several kinds of learning strategies among underachieving students. Students may be well aware of this, as a majority of underachieving gifted students in Colangelo et al.'s (1993) study stated they needed to improve their learning strategies (as opposed to just 15% of high achieving students; furthermore, this difference was somewhat specific, as no such gap appeared on questions asking whether students needed help with educational plans or with personal problems). The finding that learning strategies mediate the association between parental support and achievement is also in line with the AOM (Veas, Castejón, O'Reilly & Ziegler, 2018). Direct tests of the model have also provided support for it (Ritchotte, Matthews & Flowers, 2014; Rubenstein et al., 2012). In-depth interviews with a small sample of gifted students confirmed that the constructs in the AOM are subjectively important for these students as well (Barbier, Donsche & Verschueren, 2019). The motivational variables used in the current study do not match these components so exactly as to constitute a direct test of the AOM, but they do represent the variety of motivational components summarized in the model. Furthermore, we tested whether it is true that these components cannot compensate for one another, as this assumption of the AOM currently lacks empirical validation.³

1.3.2 Personality

We chose the Big 5 personality variables (Costa & McCrae, 1992) as another branch of predictors because there is surprisingly little empirical research on how underachievers differ (or do not differ) from other students in this major domain of interindividual differences. In general, the Big Five do not show uniform effects on academic achievement, with the exception of a small positive association with conscientiousness (O'Connor & Paunonen, 2007); since most underachievement research compares underachievers with high achievers only and thus basically looks for correlations with achievement (within the group of high-IQ students), this

³ Ritchotte et al. (2014) reported support for it, but the test they reported just reiterated the main effect of the motivational components; it did not test for compensation, which concerns the interaction between the components.

overall null correlation might have led researchers to regard the Big Five as an unattractive object of study. However, the question remains whether the general null results can be extrapolated to the particular subgroup of underachievers, especially under the framework of the present study. The findings that do exist, as listed in Reis and McCoach (2000)'s review, are diverse and contradictory; almost no studies have since helped to clarify the situation, with the exception of Preckel and colleagues (2006) who reported that underachievers displayed lower conscientiousness; they did not report testing the other Big Five dimensions.

It is often stressed that underachievers are a very heterogeneous group (Siegle, 2018). This might be one explanation for the diversity of findings on the topic, but it might also partly be due to the different empirical approaches. However, to our knowledge, McCoach and Siegle's (2003) finding of larger variances in a group of underachievers than in a group of high achievers (concerning motivational attitudes) as well as a rather qualitative finding by Smith (2007) constitute the only reported evidence of the conjured larger heterogeneity of underachievers in contrast to other groups. We therefore tested whether such a pattern held in our sample as well.

2 Materials and methods

The present study drew on a large sample collected in TwinLife, a panel study that surveys pairs of twins as well as their extended families in Germany, with a focus on following and explaining different life courses (for an overview, see Mönkediek et al., 2016). The sample is mostly representative for the German population, although families with higher socio-economic status are somewhat overrepresented (Lang & Kottwitz, 2017). Data from the first two data collection phases are available as scientific use files via the GESIS data archive (<https://doi.org/10.4232/1.13208>).

As one aspect of tracking a person's road to success in life, TwinLife includes a number of motivational and environmental variables pertaining to academic success. Some of these self-report scales were used as approximations of the AOM constructs in this study: learning valuation (meant to resemble goal valuation in the AOM); self-efficacy and academic self-concept (self-efficacy in the AOM); and parenting style (reported by the children), which consisted of scales measuring parental support, autonomy and control (environmental perceptions). TwinLife did not include a direct measure of self-regulation in the first data wave (which was used here); our proxy for successful (or unsuccessful, rather) self-regulatory behavior necessary as a further link to academic achievement according to the AOM therefore was hyperactivity (a subscale of measuring externalizing behavior).

After four joint years of elementary school, Germany's school system is stratified, with grammar schools (Gymnasium) being the highest level in the sense that the degree obtainable in these schools (Abitur) is a prerequisite for studying at a German university and the most valuable for the labour market. This creates an issue when trying to study underachievement: A student can be successful or unsuccessful in terms of better or worse grades within the school that they attend, but educational attainment in the long run is also characterized by what level of school the student attends in the first place (see Uhlig, Solga & Schupp, 2009, whose study looks at

underachievement through this alternative perspective). As the latter variable (the different school levels in the stratified system) constitutes merely an ordinal scale, a reliable metric to compare good grades in a lower-level school to bad grades in grammar school is not feasible. The analysis was therefore restricted to over- and underachievement within grammar schools. Students' grades were collected by photographing the most recent report card (these are issued twice a year). If taking a photo was not possible or not agreed to, parents were asked to provide the grades for German and Math (this was the case for only about 1% of the sample). When analyzing German school grades, it is common practice to only gather data for the subjects of German and Math, as these are the only two subjects that are taught in virtually every school. This procedure was followed in the present study as well; when report cards were not available, parents were asked to report (only) on these two subjects. All report cards dated from the same year as the general survey used for the other variables, and students were excluded from the sample if they were no longer attending school at the time of the survey (e.g., their most recent report card was their final one). With these restrictions, a total of 612 students (324 female; mean age = 14.6 years, $SD = 2.8$) remained in the sample. They came from 377 different families and included 103 complete monozygotic (MZ) pairs, 104 dizygotic (DZ) pairs, 36 twins from incomplete MZ pairs (because their co-twin did not attend a gymnasium school, for example); 75 twins from incomplete DZ pairs, and 87 siblings of twins.

General intelligence was measured using a short version of the Culture Fair Test, or CFT (Weiß & Osterland, 2012), administered via computer and consisting of four subtests measuring reasoning, figural reasoning, figural classification and matrices. It is a frequently used (s. Kuhn, Holling & Freund, 2008) test aiming to measure fluid intelligence (i. e., capability of solving problems and processing complex information, Cattell, 1963) through tasks that are designed to be as "culturally fair" as possible in that they rely on basic, universal contents such as geometric patterns, as opposed to specific knowledge that depends more strongly on an individual's culture and language. A factor analysis of the four scales yielded a clear unifactorial solution. Factor scores were used as an overall index of cognitive ability (although this did not really make a difference as they were very highly correlated with simple sum scores). Great care was taken to exclude (possibly) invalid cases by having home interviewers note any irregularities as well as by having two independent raters judge all cases with suspicious data patterns, i. e. subtests diverging strongly from one another; for further details, see Gottschling (2017).

Both variables (grades as well as cognitive test scores) were standardized by age and sex (as is recommended when using twin data; McGue & Bouchard, 1984).⁴

⁴ It was also checked whether grade averages differed substantially across Germany's 16 federal states; while statistically significant, the differences were small enough in effect size, with $\eta^2 = 0.053$, that standardizing grades by state was deemed likely to be spurious rather than helpful: some small states were only represented by a handful of students and it is unknown whether the differences are even attributable to unwelcome method variance.

The instruments used in TwinLife to measure the predictor variables in this study are as follows. Self-efficacy, learning motivation and academic self-concept were measured using existing German-language short scales (Beierlein, Kovaleva, Kemper & Rammstedt, 2012 for self-efficacy; Spinath et al., 2002, for learning motivation; Dickhäuser, Schöne, Spinath & Stiensmeier-Pelster, 2002, for academic self-concept). Cronbach's alpha coefficients in our sample were 0.71, 0.70 and 0.72, respectively. Sample items include "I can rely on my own abilities in difficult situations" for self-efficacy and "In school, I want to learn as much as possible" for learning motivation. Participants indicated how well these statements described them on a five-point scale ranging from "doesn't apply at all" to "applies exactly". For academic-self concept, children were asked to complete an item such as "I am ... for school" with options ranging from, for example, "not talented" to "very talented".

Parenting style was adapted from Spinath and Wolf (2006) and measured parental support, parental autonomy and parental control with three items each; these subscales reached reliabilities between 0.68 and 0.79. Participants indicated the extent of their agreement to statements such as "My parents console me and help me when I have problems in school" (support), "When my parents help me with my studies they encourage me to find the solution myself" (autonomy) and "When I get a poor grade, my parents complain and demand that I work harder" (control).

Hyperactivity was part of a scale on externalizing behavior adapted from Goodman et al. (1998). Cronbach's alpha for the five items was 0.72. An example item is "I am easily distracted, I find it difficult to concentrate". The Big Five personality subscales of extraversion, neuroticism, agreeableness, openness to experience and conscientiousness were assessed using the Big-Five Inventory SOEP (s. Gerlitz & Schupp, 2005). Participants were asked to indicate how strongly (on a 7-point scale in this case) they agreed to statements about themselves, all of which began "I see myself as someone who..." and then mentioned various traits typical for each of the five factors, e. g., "...worries a lot" for neuroticism or "...tends to be lazy" for (low) conscientiousness. These mini scales with 3–4 items each achieved Cronbach's alphas ranging from 0.55 to 0.74.

Data were collected through home visits by trained interviewers. Because of the somewhat sensitive topics, most of these measures were assessed through self-administered computerized testing; for learning motivation and academic self-concept, however, the interviewers asked the questions directly and recorded the answers. More detailed information, including all of the items, is available from the TwinLife Scales Manual (Baum, Klatzka, Iser & Hahn, 2020).

As explained in the introductory section, students were classified as expected low and high achievers or as unexpected low and high achievers (i.e., under- and overachievers) by selecting those students who combined a notably above- or below-average IQ with above- or below-average grades. Of course, specifying the cut-off for "notably" above- or below-average scores is an arbitrary choice; there is a tradeoff between the sample size that is retained after the selection and the validity and significance of the cases. Too lax a criterion will result in more spurious cases due to error variance, as well as cases that can technically be termed, for example, underachievement, but do not actually represent severe difficulties in school (see Sect. 1.2). Cutting off both the standardized intelligence scores and the standardized

grade averages derived earlier at Z values of 0.5 above and below the mean resulted in a reasonable balance between sample size and validity concerns (see below). Additionally, although defining cutoffs is in principle an arbitrary decision, school grades do offer some reference points. Grades in the German school system range from 1 to 6, with 1 being the best grade. An important cutoff lies between 4 and 5, as a grade of 4 or better "is generally required in each of the subjects that have a bearing on promotion" (Lohmar & Eckhardt, 2014). Therefore, there can be little doubt that a grade of 5 or 6 would definitely fall below a reasonable cutoff for low achievement. However, these grades are so rare (about 2% of the cases in our sample) that selecting only students with a 5 or 6 in either German or Math would have reduced the sample size drastically. In fact, the distribution of grades is skewed towards good grades so strongly that a few students were classified as underachievers or low achievers with a cutoff of 0.5 SD while receiving a 3 ("satisfactory") in both subjects, despite the fact that this selection was already much more restrictive than the usual regression approach. These cases were excluded in order to ensure that all low achievers and underachievers were at least on the brink of failing a class. Despite our already relatively strict selection approach, there were still 9 such cases out of originally 30 underachievers and 19 out of 82 low achievers. As far as our German sample can be generalized, and as far as readers follow our interpretation of „serious“ underachievement, it thus appears to be a remarkably rare phenomenon: In total, there were 352 different students who fell either into the high IQ group or the low achievement group, or both; the latter case, the overlap between both categories, represents underachievement, but happened only in 21 cases out of the 352.

In total, 21 underachievers, 30 overachievers, 63 low achievers and 113 high achievers remained after these selection procedures (the uneven distribution of low and high achievers is partly due to the exclusion of some low achievers as mentioned above and partly due to the categorical nature of the achievement variable). As the sample from which these students were selected included 612 students, 63% were not classified into any of the extreme groups as either their IQ or their grades, or both, were more or less average. A total of 3% and 5% were identified as underachievers or overachievers, respectively. Note that the usual approach of using a difference of 1 SD between standardized grades and IQ scores results in about 15% selected cases in either direction, but, in our view, will include many cases whose classification as practically meaningful underachievement is doubtful. That our approach was comparatively strict can also be seen from the fact that underachievers represented about 10% of high-IQ students in our sample; in a literature review by White et al. (2018), this figure ranged from 9 to 28% in different studies (with different methodologies and samples).

The 227 students came from 182 different families and included 17 complete MZ pairs and 20 complete DZ pairs. No family included two underachievers, while three families included two overachievers each (MZ twin pairs in two cases, siblings in one case). The average age was 14.6 years ($SD=2.8$); due to the nature of the panel study, which consisted of cohorts separated by about 6 years, a majority of students were either 11 years old (99 cases) or 17 years old (72 cases).

Among the predictor variables used in the discriminant analysis, 2.4% of the values were missing. Little's MCAR test, a conservative test assessing whether data

are missing completely at random (Little, 1988), was not significant ($\chi^2 = 124.13$, $df = 104$, $p = 0.09$), indicating that it was sufficient to replace missing data using the Expectation Maximization procedure rather than using the more complex approach of multiple imputation (Tabachnick & Fidell, 2001).

3 Results

Table 2 displays zero-order correlations between the predictor variables within the final sample, i. e. those students that were classified as either underachievers, overachievers, or belonging to a suitable control group; descriptive statistics are shown in Table 3, grouped by achievement category. Results of the discriminant analysis are shown in Tables 4, 5.

Box's M test was insignificant (Box's $M = 286.8$, $p = 0.357$), indicating that the assumption of homogeneity of covariances necessary for discriminant analysis was not violated. The first discriminant function explained 66% of the variance (canonical $R^2 = 0.26$), the second function explained 32% (canonical $R^2 = 0.14$); the contribution of the third function (3% of the variance, canonical $R^2 = 0.01$) was insignificant. Wilks's lambda for all functions combined was $\Lambda = 0.63$, $X^2(36) = 101.4$, $p < 0.001$; over and above the first function, the other two still differentiated the groups significantly, with $\Lambda = 0.85$, $X^2(22) = 36.8$, $p = 0.025$.

The position of the group centroids on the first function suggest that it almost exclusively separated successful students from unsuccessful ones: Underachievers (group centroid at -0.79) were grouped together with low achievers (-0.75), while overachievers (0.44) and high achievers (0.45) were indistinguishable at the other end. In other words, as far as underachievers are concerned, this function highlighted those variables that differentiated underachievers from successful students, regardless of whether those are high achievers or overachievers.

The canonical correlations indicate that the successful students reported a higher academic self-concept, higher self-efficacy, learning motivation and conscientiousness; they reported less hyperactivity and a more functional style of parental involvement (more autonomy and support, less control). Note that discriminant analysis yields two different outputs; apart from the canonical correlations, the weights in the discriminant functions themselves indicate each variable's unique contribution to the empirical separation of the groups. In this regard, not all of the variables in which the two groups differed were crucial for distinguishing them (for example, if all the other variables are known, hyperactivity has little incremental value beyond them). Academic self-concept was the most prominent separating variable.

The group centroids of the second function indicate that, after separating successful and unsuccessful students, classification can mostly be improved by contrasting low-IQ and high-students, and in particular by contrasting underachieving students (group centroid at -0.73) to overachieving students (0.77). Expected low (0.25) and high achievers (-0.21) were separated to a lesser degree, but were more similar to the students who were in the same IQ group as they were.

In other words, this function differentiates whether a student over- or underperforms in terms of IQ, and it indicates which aspects are associated with

Table 2 Zero-order correlations for the predictor variables in the different achievement groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Parental support											
(2) Parental autonomy	.79**										
(3) Parental control	-.30**	-.22*									
(4) Learning motivation	.22*	.25*	-.17*								
(5) Self-efficacy	.16*	.23*	-.17*	.40**							
(6) Academic self-concept	.21*	.17*	.08	.21*	.26**						
(7) Hyperactivity	-.10	-.13	.32**	-.23*	-.27**	-.27**					
(8) Neuroticism	-.09	-.14	-.10	.11	-.21*	.03	.01				
(9) Extraversion	.26**	-.16*	-.06	.06	.23*	.18*	.03	-.21*			
(10) Conscientiousness	.21*	.22*	-.31**	.33**	.34**	.29**	-.44**	.12	-.02		
(11) Agreeableness	.28**	.25**	.19*	.24**	.04	.21*	-.23*	.20*	.06	.29**	
(12) Openness	.14*	.06	-.07	.27**	.13	.09	-.12	-.02	.24**	.12	.23*

* $p < 0.05$. ** $p < 0.01$

Table 3 Group means and standard deviations for the predictor variables in the different achievement groups

	Underachievers (<i>n</i> = 21)	Low achievers (<i>n</i> = 63)	High achievers (<i>n</i> = 113)	Overachievers (<i>n</i> = 30)
Parental support	9.24 (2.12)	8.86 (2.74)	9.75 (2.27)	9.61 (2.71)
Parental autonomy	8.71 (2.63)	9.03 (2.61)	9.53 (2.41)	10.05 (2.37)
Parental control	4.71 (2.12)	4.79 (3.22)	2.99 (3.03)	2.73 (2.75)
Learning motivation	8.81 (1.83)	8.84 (1.88)	9.54 (1.61)	10.14 (1.63)
Self-efficacy	7.86 (1.93)	8.37 (2.08)	9.18 (1.50)	9.47 (1.76)
Academic self-concept	6.58 (0.67)	6.83 (1.33)	7.52 (1.06)	7.57 (0.77)
Hyperactivity	3.71 (2.05)	3.88 (2.41)	2.58 (1.93)	3.05 (2.18)
Neuroticism	11.86 (4.33)	11.71 (5.11)	11.17 (4.68)	11.17 (4.01)
Extraversion	10.57 (5.04)	12.22 (4.20)	11.27 (3.67)	12.27 (3.62)
Conscientiousness	10.00 (2.92)	11.16 (2.96)	12.74 (2.82)	13.43 (3.17)
Agreeableness	13.00 (2.05)	13.56 (3.44)	13.69 (2.89)	14.57 (2.53)
Openness	17.19 (4.26)	16.06 (4.06)	17.65 (3.97)	16.43 (3.97)

Variable ranges were: 0–12 for the parental involvement variables, learning motivation, self-efficacy, and academic self-concept; 0–10 for hyperactivity; 0–18 for E, C and A, and 0–24 for N and O

Table 4 Values of the functions at group centroids for the two first discriminant functions

	Function 1	Function 2
Underachievers	– .79	– .73
Low achievers	– .75	.25
High achievers	.45	– .21
Overachievers	.44	.77

The third discriminant function is omitted because of its lack of significance

Table 5 Canonical correlations of each predictor with the first two discriminant functions as well as (in brackets) weights in the discriminant functions

	Function 1	Function 2
Parental support	.25 (.06)	– .10 (– .21)
Parental autonomy	.23 (– .04)	.18 (.10)
Parental control	– .52 (– .48)	– .04 (.30)
Learning motivation	.40 (.15)	.22 (– .87)
Self-efficacy	.48 (.22)	.22 (.29)
Academic self-concept	.58 (.57)	.12 (.19)
Hyperactivity	– .47 (– .03)	– .22 (– .25)
Neuroticism	– .10 (– .26)	– .01 (.20)
Extraversion	– .07 (– .38)	.33 (.34)
Conscientiousness	.58 (.28)	.31 (.37)
Agreeableness	.13 (– .17)	.26 (– .20)
Openness	.22 (.21)	– .30 (.14)

underachievement (and its mirror image of overachievement) rather than just with low achievement in general (as that part of the variance has already been dealt with in the first function).

The explanatory value of this function was, while significant, rather small (explaining 11% of the overall variance). The most substantial canonical correlations were present among the Big Five variables, with underachievers reporting lower (and overachievers higher) extraversion and conscientiousness, as well as higher openness to experience. Motivational variables generally pointed into the direction of a more dysfunctional pattern in underachievers (lower motivation and belief in themselves, higher hyperactivity, less parental support), albeit with only small coefficients.

The group means (see Table 3) offer tentative support for interpreting the results as a meaningful pattern rather than random variance, as the group differences emerged about equally in both directions in all of the notably contributing variables – overachievers mirrored underachievers in their differences from the comparison groups.

Using the discriminant functions, 45.8% percent of the cases could be classified correctly (as opposed to 25% by chance alone). However, this value dropped to 36.1% when each case was excluded from predicting itself (leave-one-out cross-classification; Molinaro et al., 2005) pointing to some likely overfitting in the results. Using a pre-defined group size is perhaps a more realistic approach, as the expected proportion of underachievers under a given selection criteria is known by definition; on the other hand, this means that mostly high achievers and low achievers are predicted, not the more „interesting“ cases of under- or overachievement. This approach yielded a correct classification of 59.0% without and 53.3% with cross-classification.

Based on the variables used for classifications, underachievers appeared to resemble (non-underachieving) low achievers (their likelihood of belonging to that group was estimated at 26% on average) somewhat more closely than high achievers (20%). This analysis was repeated with academic self-concept excluded because it could be argued that underachievers resemble other low achievers simply because they have basically stated (indirectly) that they are low achievers (although it should be noted that academic self-concept and grades correlated only moderately at $r=0.36$). The difference shrunk to 25% vs. 23%. A similar yet reverse pattern was apparent for overachievers, who were more similar to expected high achievers (with comparable grades) than to expected low achievers (with comparable IQ) with values of 26% vs. 20% when including and 25% vs. 21% when excluding academic self-concept. Another way of operationalizing the similarity of the groups is to calculate the distance between the group centroids; at 0.97, the distance between underachievers and low achievers was smaller than between underachievers and high achievers (1.34).

In order to test the hypothesis that underachievers are more heterogeneous as a group than the other “types” of students, a Levene test was performed on all variables as well as the discriminant scores for both functions. Variances for the overall discriminant scores did not differ significantly; among the individual variables, self-efficacy, academic self-concept, agreeableness and extraversion displayed heterogeneous variances across groups (not accounting for multiple testing). However, descriptive values (see Table 3) show that instead of underachievers, it is the group

of other “expected” low achievers that displayed the highest variance in these variables, except for extraversion. In academic self-concept and agreeableness, underachievers actually displayed the lowest variance among all four groups. Therefore, we found no evidence to support the assumption that there is more variation among underachievers in the motivational and personality variables examined here – if anything, the contrary is true in our sample.

The assumption made in the AOM that the motivational precursors to achievement cannot be compensated for among each other was tested by checking whether the minimum among these (standardized) variables, as well as the number of variables below a certain low threshold (1 SD below the mean, approximately following Matthews et al., 2014), has any incremental validity in predicting underachievement when the mean level of these variables is taken into account. Underachievement was operationalized via the regression approach (see introduction) in this instance.⁵ As expected, the mean Z-scores in the variables that most closely resembled the AOM components (learning motivation, self-efficacy, academic self-concept and parental support as the closest proxy to environmental support) correlated substantially with the underachievement score ($r=0.37$, $p<0.001$; the coefficient is positive because higher scores in the regression approach denote overachievement). Their minimum value as well as the number of components below the threshold correlated similarly highly, and in plausible directions, with the underachievement score ($r=0.27$ and $r=-0.37$, respectively; p 's <0.001). However, when the mean was introduced as a control variable in a partial correlation, both associations became insignificant ($r=-0.09$, $p=0.17$ for the minimum value and $r=-0.11$, $p=0.08$ for the number of low scores). Thus, the non-compensation hypothesis was not supported.

4 Discussion

4.1 General discussion

The present study approached the issue of underachieving students under a slightly extended framework than previous research. Motivational variables similar to those in the Achievement Orientation Model as well as the Big Five personality variables were examined as traits that might differentiate underachievers from other students. This was in part a replication of established differences between underachievers and high achievers; adding the simultaneous comparison to low achievers (and overachievers) as well as, in terms of variables used, the Big Five, were the novel parts of the study. The first discriminant function, separating both kinds of successful students from both kinds of unsuccessful ones, indeed reiterates the point that motivational variables are strongly associated with academic achievement. Correlations between achievement and a functional parenting style were also replicated. The Achievement Orientation Model as a framework for understanding what can

⁵ This was done in order to test as liberally as possible (by not discarding any variance in the data) so that a possible null result might not be explainable by too conservative a test.

get in the way of a gifted student translating their cognitive abilities into educational attainment thus receives some support from our findings.

The second point is somewhat less clear; while the second function in the discriminant analysis significantly improved classification, coefficients were generally small (and possibly unstable, see below).

However, with these restrictions in mind, our study adds to previous research in justifying the construct itself: It appears that underachievers do in fact differ not only from other highly intelligent students, as previous research had already demonstrated, but that some differences to other low achievers are also discernible, albeit smaller. Yet this difference is only half the story: the discriminant analysis included four groups, and overachievers appeared to be more or less a mirror image of underachievers. This corroborates a so far very limited research body on overachievement (Sparfeldt et al., 2010). The bottom line is that those relatively rare cases where cognitive abilities—as the most important predictor of academic achievement—are “overruled” by other factors are not solely the product of random (error) variance; those cases instead appear to have something in common.

According to our results, these things in common are, among others, personality differences. It is a somewhat novel finding that the specific discrimination between underachievers and low achievers (as well as overachievers and high achievers; a parallel we will not mention explicitly hereafter for the sake of brevity) was in part due to the Big Five factors. Low agreeableness, low conscientiousness and low extraversion were typical for our sample of strongly underachieving students; further research might corroborate or disprove these tentative findings. The finding that the underachievers in our sample tended to be introverted could help explain the findings cited above (Rost & Hanses, 1997, and Siegle et al., 2020) which indicate that teachers often have trouble recognizing which of their students underachieve, and why.

The motivational variables exerted less influence in discriminating underachievers and low achievers. Although the variables mostly pointed in the expected direction numerically, it cannot be denied that academic self-concept in particular was not very useful for identifying underachievers beyond the function that separated the successful from the unsuccessful groups. The finding that underachievers scored lower on academic self-concept than successful students is not surprising as this merely reflects the obvious facts on their report cards; the plausible conjecture that their self-concept might be even lower than that of other low achievers (because they might compare it to a higher standard, suspecting that they “should” be able to do better) could not be confirmed. (Note that, as the descriptive values in Table 3 show, this was not due to a floor effect.) Low scores on the more generally phrased items assessing self-efficacy were somewhat more typical for underachievers, but only slightly. When trying to identify students who have “hidden potential” that has not been brought to life yet, it might seem reasonable to think that the most disheartened students may be those who are falling short of their potential because they just are not believing in themselves. However, not believing in their capabilities seems to be almost as typical for any unsuccessful student as for underachievers. In this light, it is unsurprising that an intervention targeting this aspect was unsuccessful (Obergrösser & Stoeger, 2015) as it might have tackled a symptom rather than a cause.

Note that this does not necessarily mean that underachievers do not know they are underachieving. Students were asked about their practical ability to succeed, not about their intelligence (or the relation between these two constructs).

Besides not being identical to either other low achievers or other highly intelligent students, detailed results of the discriminant analysis indicated that in terms of the motivational and personality variables used here, underachievers are hardly even closer to one group than the other – they cannot clearly be put into either box. They are somewhere in between, a result echoing findings by Hanses and Rost (1998) in one of the few other studies that included similar comparisons.

Although our results do provide some indirect support the AOM, we found no evidence for its assumption that if one of the different facets of a student's motivation is particularly low, they are more likely to underachieve than a student with the same mean level of motivation, but a more balanced distribution. That being said, the variables we used only approximate those in the AOM.

In the sample used in this study, there was also no indication for another suggestion that can frequently be found in the literature on underachievement—that they are more heterogeneous than other groups of students. If anything, our results point to the contrary. Further empirical clarification might be warranted, especially because the way we tested it, using variance as a proxy for heterogeneity, certainly does not answer the question once and for all. It does not follow, of course, that underachievement is a uniform phenomenon.

4.2 Limitations and Strengths

One obvious caveat of the present study is that no causal inference can be made; for example, it remains unclear whether an unhelpful manner of parental involvement actually causes students' difficulties—parents might also simply be reacting to their children's bad grades.

The results also warrant some replicability concerns. In balancing the trade-off between sample size and robustness of the selection, we focused on validity, which necessarily diminished the potentially large sample size to only 51 overachievers and underachievers in total. With ten predictors in the discriminant analysis, a sample size of 21 in the smallest group can only offer tentative results, even though the comparison groups were larger. The fact that classification worsened notably when leave-one-out classification was used suggests that some non-negligible overfitting was indeed taking place. Less strict inclusion criteria would have retained a larger part of the original sample and therefore would have allowed for more reliable conclusions, but their validity would have been doubtful. Part of the problem was the fact that bad grades (4–6 out of the 1–6 range in the German school system) were quite uncommon in the sample, perhaps reflecting a general tendency of grade “inflation” in Germany (Grözinger & Baillet, 2015). Consequently, students with clear-cut underachievement are actually relatively rare. An argument could be made that underachievement can also entail achieving mediocre grades despite great potential—especially as the aforementioned inflation slowly changes the relative meaning of these grades. But even if

we had followed this logic, including these cases in our analysis would have run the risk of grouping very different cases together under one label and thereby comparing apples and oranges. Besides the sample size, it must also be noted that due to the nature of the panel study, short scales were used to measure the predictor variables; more reliable instruments might yield more stable results.

The fact that we used a sample from a twin study also warrants some caution in interpreting the results as the underlying observations are not fully independent of each other (Carlin et al., 2005). In order to maintain a reasonable sample size, we decided against including only one randomly selected member of each twin pair of family. The sample used for the final analysis ended up including 20 MZ and 17 DZ pairs; still, due to the selection process of filtering out students who did not fit into any of the achievement groups we were interested in, the majority (60%) were students who were the sole member of their family to be included. Note also that the 21 underachievers actually came from 21 different families—if their twin or sibling was part of the sample, they were in other groups. Among the 30 overachievers, there were two MZ pairs and one sibling pair. Overall, this does not give the impression that the differences between the groups could be explained by artificially increased similarity within them. On a sidenote, the fact that the underachievement and overachievement groups do *not* consist of notably many (MZ) pairs could be taken as a tentative indication that the heritability of over- or underachieving, as opposed to academic achievement per se, is not strikingly high. Exploring this issue was not an aim of this paper, but could be a question for future research.

On the other hand, it is notable that the study yielded meaningful results even though its range was necessarily restricted because it was limited to grammar school students. When a wider perspective is taken, other factors might come into play as well. As shown by Uhlig et al. (2009), socio-economical background is important (at least in Germany) in determining students' overall academic course. This factor is beyond the students' control, but motivation and personality—besides cognitive abilities, of course—then seem to “finetune” how successful they are within the academic setting they find themselves in.

It is noteworthy that the items used in this study operationalize their target constructs, such as motivation, on a very global level, without investigating root causes. We measured whether a child was motivated or not, and what grades were achieved, but do not know which external influences might have led up to this current state. In particular, part of the existing literature on underachievement indicates that peers and teachers can play an important role in shaping students' educational paths (e. g., Baker, 1998; Woodward & Fergusson, 2000; Boehnke, 2008; see also Sect. 1.3). A striking example comes from Fong and Krause (2014), who found that underachieving college students did not necessarily lack academic self-efficacy, but reported far less encouragement (“social persuasions”) from others, e. g. from instructors, than achieving students. Asbury et al. (2016) identified monozygotic twin pairs whose academic achievement differed and asked both them and their parents about possible reasons; different teachers were among the most frequent causal attributions. The current study does not allow conclusions as to why the different groups in our

sample differ the way they do and what role teachers and peers may have played in shaping, or reinforcing, these differences.

As a final point, some criticisms directed at the construct of underachievement, or its interpretation, are worth mentioning. On a very basic level, it should be remembered that underachievers underachieve only relative to their IQ. Stating that they “do not fulfill their potential” assumes that IQ is an all-encompassing, precise measure of a person’s potential (Ziegler et al., 2012). Instead, IQ as a predictor of achievement is just that: one predictor. A meteorologist who is surprised by a rainy day after forecasting sunshine will not claim that the day “did not fulfill its potential”, but look for factors that explain the deviation. In the same sense, achievement scores not lining up perfectly with IQ can only come about by a) measurement error or b) other influences on achievement not accounted for. The second possibility is not only more promising, but has also found support through previous research as well as our results. Also note that within the underachievement paradigm, any factor that prevents a student from doing well in school *and on an IQ test* will, by definition, not be uncovered.

It has also been pointed out that underachievers do not necessarily have to be unhappy about the situation; they might have decided for themselves that other things than school are more important to them and be fine with that choice (Figg, et al., 2012). Furthermore, the fact that underachievers as a group did not appear to be unusually heterogenous with respect to the variables used here should not be misinterpreted to mean that there is a one-fits-all recipe to help them translate their cognitive abilities into achievement. Individual perspectives are necessary and could be a focus of future research.

5 Conclusions

The present study solidified the empirical basis on which researchers and practitioners alike refer to students with low grades despite high cognitive abilities as a distinct group. That they differ from other highly intelligent students in various personal and environmental aspects was already established; besides replicating these differences, the patterns found in our sample of 51 under- or overachievers and 176 comparison students suggest that underachievers also differ from fellow unsuccessful students in terms of their personality. Motivational variables play a smaller role (if any); while motivation and parental support are lower among underachievers than among their achieving counterparts, this deficit is not extraordinary compared to other low achievers. Explaining the specific phenomenon of underachievement, as opposed to low achievement in general, through motivational problems therefore might miss the point; this is somewhat at odds with previous research and calls for future studies clarifying this issue. A particular goal of the current study was to strengthen conclusions about underachieving students as a group, both by defining relatively strict criteria for identifying underachievement in the study sample as well as by comparing them to relevant groups beyond other gifted students alone. In this way, we believe our study adds to the existing literature on underachievement by utilizing a more nuanced methodological approach.

In practical terms, this research suggests that while targeting motivation in unsuccessful students is not necessarily wrong, it is not specifically underachievers who are in need of such an intervention. The differences in the Big Five variables do not immediately suggest practical steps, but one point of interest could be the unusually low conscientiousness reported by underachievers; rather than lacking motivation, some of the might rather be in need of being taught skills and strategies to help them organize their academic endeavors effectively.

Acknowledgements This research was supported by a grant from the German Research Foundation awarded to Martin Diewald (DI 759/11-1), Rainer Riemann (RI 595/8-1), and Frank M. Spinath (SP 610/6-1). The TwinLife project received ethical approval from the German Psychological Association (protocol numbers: RR 11.2009 and RR 09.2013).

Funding Open Access funding enabled and organized by Projekt DEAL.

Availability of data and material Data from the TwinLife panel study are available as scientific use files via the GESIS data archive (<https://doi.org/10.4232/1.13208>).

Code availability Statistical analyses were conducted using SPSS, Version 25 (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.)

Declarations

Conflict of interests The authors declare that they do not have any conflict of interest.

Ethical statement The TwinLife project received ethical approval from the German Psychological Association (protocol numbers: RR 11.2009 and RR 09.2013).

Informed consent Informed consent was obtained from all individual participants included in the study or from their legal guardians.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Asbury, K., Moran, N., & Plomin, R. (2016). Nonshared environmental influences on academic achievement at age 16: A qualitative hypothesis-generating monozygotic-twin differences study. *AERA Open*, 2(4), 2332858416673596.
- Baker, J. A., Bridger, R., & Evans, K. (1998). Models of underachievement among gifted preadolescents: The role of personal, family, and school factors. *Gifted Child Quarterly*, 42(1), 5–15. <https://doi.org/10.1177/001698629804200102>
- Balduf, M. (2009). Underachievement among college students. *Journal of advanced academics*, 20(2), 274–294.

- Barbier, K., Donche, V., & Verschueren, K. (2019). Academic (under) achievement of intellectually gifted students in the transition between primary and secondary education: An individual learner perspective. *Frontiers in Psychology, 10*, 2533. <https://doi.org/10.3389/fpsyg.2019.02533>
- Baum, M.A., Klatzka, C.H., Iser, J. & Hahn, E. (2020). TwinLife scales manual: F2F1, CATI 1, & F2F2. *TwinLife Technical Report Series*, 08. ISSN 2512–403X. Bielefeld University / Saarland University.
- Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2012). Ein Messinstrument zur Erfassung subjektiver Kompetenzerwartungen: Allgemeine Selbstwirksamkeit Kurzskala (ASKU) [An instrument to measure subjective competency expectations: The general self-efficacy scale (ASKU)]. (GESIS-Working Papers, 2012/17). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften.
- Boehnke, K. (2008). Peer pressure: A cause of scholastic underachievement? A cross-cultural study of mathematical achievement among German, Canadian, and Israeli middle school students. *Social Psychology of Education, 11*(2), 149–160. <https://doi.org/10.1007/s11218-007-9041-z>
- Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R., & Dwyer, T. (2005). Regression models for twin studies: A critical review. *International Journal of Epidemiology, 34*(5), 1089–1099. <https://doi.org/10.1093/ije/dyi153>
- Castejón, J. L., Gilar, R., Veas, A., & Miñano, P. (2016). Differences in learning strategies, goal orientations, and self-concept between overachieving, normal-achieving, and underachieving secondary students. *Frontiers in psychology, 7*, 1438.
- Colangelo, N., Kerr, B., Christensen, P., & Maxey, J. (1993). A comparison of gifted underachievers and gifted high achievers. *Gifted Child Quarterly, 37*(4), 155–160.
- Cone, T. E., & Wilson, L. R. (1981). Quantifying a severe discrepancy: A critical analysis. *Learning Disability Quarterly, 4*(4), 359–371. <https://doi.org/10.2307/1510737>
- Costa, P. T., Jr., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653–665. [https://doi.org/10.1016/0191-8869\(92\)90236-I](https://doi.org/10.1016/0191-8869(92)90236-I)
- Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept Konstruktion und Überprüfung eines neuen Instrumentes [The Academic Self Concept Scales Construction and Evaluation of a new Instrument]. *Zeitschrift für differentielle und diagnostische Psychologie ZDDP, 23*(4), 393–405.
- Dixon, R. M., Craven, R., & Martin, A. (2006). Underachievement in a whole city cohort of academically gifted children: What does it look like? *Australasian Journal of Gifted Education, 15*(2), 9–15.
- Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology, 104*(2), 439–451. <https://doi.org/10.1037/a0026280>
- Figg, S. D., Rogers, K. B., McCormick, J., & Low, R. (2012). Differentiating low performance of the gifted learner: Achieving, underachieving, and selective consuming students. *Journal of Advanced Academics, 23*(1), 53–71.
- Gerlitz, J. Y., & Schupp, J. (2005). Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP [On measuring personality using the Big Five in SOEP]. *DIW Research Notes, 4*, 2005.
- Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry, 7*(3), 125–130. <https://doi.org/10.1007/s007870050057>
- Gottschling, J. (2017). Documentation TwinLife Data: Cognitive Abilities. *TwinLife Technical Report Series. Bielefeld/Saarbrücken*. <https://dbk.gesis.org/DBKSearch/download.asp?db=D&id=60848>.
- Grözinger, G., & Baillet, F. (2015). Gibt es auch beim Abitur eine Noteninflation? Zur Entwicklung der Abiturnoten als Hochschulzugangsberechtigung – Eine Darstellung und Analyse aus Soziologischer Perspektive. *Bildung und Erziehung, 68*(4), 473. <https://doi.org/10.7788/bue-2015-0407>
- Hanes, P., & Rost, D. H. (1998). Das „Drama der hochbegabten Underachiever“: „gewöhnliche“ oder „außergewöhnliche“ Underachiever? [The “drama” of the gifted underachievers - “ordinary” or “extraordinary” underachievers?]. *Zeitschrift für Pädagogische Psychologie, 12*(1), 53–71.
- Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*(1), 127.
- Hoover-Schultz, B. (2005). Gifted underachievement: Oxymoron or educational enigma? *Gifted Child Today, 28*(2), 46–49.
- Hwang, M. H., Lee, D., Lim, H. J., Seon, H. Y., Hutchison, B., & Pope, M. (2014). Academic underachievement and recovery: Student perspectives on effective career interventions. *The Career Development Quarterly, 62*(1), 81–94. <https://doi.org/10.1002/j.2161-0045.2014.00072.x>

- Johnson, W., McGue, M., & Iacono, W. G. (2006). Genetic and environmental influences on academic achievement trajectories during adolescence. *Developmental Psychology, 42*(3), 514–532. <https://doi.org/10.1037/0012-1649.42.3.514>
- Kuhn, J. T., Holling, H., & Freund, P. A. (2008). Begabungsdagnostik mit dem Grundintelligenztest (CFT 20-R) Psychometrische Eigenschaften und Messäquivalenz [Assessing giftedness with the CFT 20-R Psychometric properties and measurement invariance]. *Diagnostica, 54*(4), 184–192.
- Landis, R. N., & Reschly, A. L. (2013). Reexamining gifted underachievement and dropout through the lens of student engagement. *Journal for the Education of the Gifted, 36*(2), 220–249.
- Lang, V., & Kottwitz, A. (2017). The sampling design and socio-demographic structure of the first wave of the TwinLife panel study: a comparison with the Microcensus. *TwinLife Technical Report Series* (Vol. 3).
- Lau, K. L., & Chan, D. W. (2001). Identification of underachievers in Hong Kong: Do different methods select different underachievers? *Educational Studies, 27*(2), 187–200. <https://doi.org/10.1080/03055690120050419>
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202.
- Lohmar, B., & Eckhardt, T. (Eds.). (2014). *The Education System in the Federal Republic of Germany 2012/2013*. Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany. Retrieved from: http://www.istp2016.org/fileadmin/Redaktion/Dokumente/documentation/dossier_en_ebook.pdf
- McCall, R. B. (1994). Academic underachievers. *Current Directions in Psychological Science, 3*(1), 15–19.
- McCall, R. B., Beach, S. R., & Lau, S. (2000). The nature and correlates of underachievement among elementary schoolchildren in Hong Kong. *Child Development, 71*(3), 785–801. <https://doi.org/10.1111/1467-8624.00185>
- McCoach, D. B., & Siegle, D. (2003). Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gifted child quarterly, 47*(2), 144–154.
- McGue, M., & Bouchard, T. J. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics, 14*(4), 325–343. <https://doi.org/10.1007/BF01080045>
- Mofield, E., & Parker Peters, M. (2019). Understanding Underachievement: Mindset, Perfectionism, and Achievement Attitudes Among Gifted Students. *Journal for the Education of the Gifted, 42*(2), 107–134.
- Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: A comparison of resampling methods. *Bioinformatics, 21*(15), 3301–3307. <https://doi.org/10.1093/bioinformatics/bti499>
- Mönkediek, B., Lang, V., Weigel, L., Baum, M. A., Eifler, E. F., Hahn, E., Hufer, A., Klatzka, C. H., Kottwitz, A., Krell, K., Nikstat, A., Diewald, M., Riemann, R., & Spinath, F. M. (2019). The German Twin Family Panel (TwinLife). *Twin Research and Human Genetics, 22*, 540–547. <https://doi.org/10.1017/thg.2019.63>
- O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971–990. <https://doi.org/10.1016/j.paid.2007.03.017>
- Obergriesser, S., & Stoeger, H. (2015). The role of emotions, motivation, and learning behavior in underachievement and results of an intervention. *High Ability Studies, 26*(1), 167–190. <https://doi.org/10.1080/13598139.2015.1043003>
- Peterson, J. S., & Colangelo, N. (1996). Gifted achievers and underachievers: A comparison of patterns found in school files. *Journal of Counseling & Development, 74*(4), 399–407. <https://doi.org/10.1002/j.1556-6676.1996.tb01886.x>
- Plewis, I. (1991). Underachievement: A case of conceptual confusion. *British Educational Research Journal, 17*(4), 377–385. <https://doi.org/10.1080/0141192910170407>
- Preckel, F., Holling, H., & Vock, M. (2006). Academic underachievement: Relationship with cognitive motivation, achievement motivation, and conscientiousness. *Psychology in the Schools, 43*(3), 401–411. <https://doi.org/10.1002/pits.20154>
- Reis, S. M., & McCoach, D. B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted child quarterly, 44*(3), 152–170.
- Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of Personality and Social Psychology, 111*(5), 780–789. <https://doi.org/10.1037/pspp0000089>

- Ritchotte, J. A., Matthews, M. S., & Flowers, C. P. (2014). The validity of the achievement-orientation model for gifted middle school students: An exploratory study. *Gifted Child Quarterly*, *58*(3), 183–198.
- 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*(2), 167–177.
- 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. <https://doi.org/10.1016/j.intell.2015.09.002>
- Rubenstein, L. D., Siegle, D., Reis, S. M., McCoach, D. B., & Burton, M. G. (2012). A complex quest: The development and research of underachievement interventions for gifted students. *Psychology in the Schools*, *49*(7), 678–694. <https://doi.org/10.1002/pits.21620>
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. <https://doi.org/10.1037/bul0000098>
- Siegle, D. (2018). Understanding underachievement. In S. I. Pfeiffer (Ed.), *Handbook of Giftedness in Children* (pp. 285–297). Springer.
- Siegle, D., McCoach, D. B., & Roberts, A. (2017). Why I believe I achieve determines whether I achieve. *High Ability Studies*, *28*(1), 59–72. <https://doi.org/10.1080/13598139.2017.1302873>
- Siegle, D., DaVia Rubenstein, L., & McCoach, D. B. (2020). Do you know what I'm thinking? A Comparison of Teacher and Parent Perspectives of Underachieving Gifted Students' Attitudes. *Psychology in the Schools*. <https://doi.org/10.1002/pits.22345>
- Smith, E. (2007). Considering the experiences of 'underachieving' and 'overachieving' students. *International Journal of Research & Method in Education*, *30*(1), 19–32. <https://doi.org/10.1080/17437270701207702>
- Snyder, K. E., & Linnenbrink-Garcia, L. (2013). A developmental, person-centered approach to exploring multiple motivational pathways in gifted underachievement. *Educational Psychologist*, *48*(4), 209–228. <https://doi.org/10.1080/00461520.2013.835597>
- Sparfeldt, J. R., Schilling, S. R., & Rost, D. H. (2006). Hochbegabte Underachiever als Jugendliche und junge Erwachsene: Des Dramas zweiter Akt? [Gifted Underachievers as Adolescents and Young Adults: Second Part of the „Drama of Gifted Underachievers“?]. *Zeitschrift für Pädagogische Psychologie*, *20*(3), 213–224. <https://doi.org/10.1024/1010-0652.20.3.213>
- Sparfeldt, J. R., Buch, S. R., & Rost, D. H. (2010). Klassenprimus bei durchschnittlicher Intelligenz – Overachiever auf dem Gymnasium [Excellent Scholastic Achievement with Average Intelligence – Overachievement in Secondary School]. *Zeitschrift für pädagogische Psychologie*, *24*(2), 147–155. <https://doi.org/10.1024/1010-0652/a000012>
- Spinath, F. M., & Wolf, H. (2006). CoSMoS and TwinPaW: Initial report on two new German twin studies. *Twin Research and Human Genetics*, *9*(6), 787–790. <https://doi.org/10.1375/twin.9.6.787>
- Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2002). *Skalen zur Erfassung der Lern- und Leistungsmotivation: SELLMO [Scales for the measurement of learning motivation and achievement motivation: SELLMO]*. Hogrefe.
- Stankov, L., & Lee, J. (2014). Quest for the best non-cognitive predictor of academic achievement. *Educational Psychology*, *34*(1), 1–8. <https://doi.org/10.1080/01443410.2013.858908>
- Steenbergen-Hu, S., Olszewski-Kubilius, P., & Calvert, E. (2020). The effectiveness of current interventions to reverse the underachievement of gifted students: Findings of a meta-analysis and systematic review. *Gifted Child Quarterly*, *64*(2), 132–165. <https://doi.org/10.1177/0016986220908601>
- Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, *47*, 1–41.
- Tabachnick, B. G., & Fidell, L. S. (2001). Cleaning up your act: Screening data prior to analysis. *Using Multivariate Statistics*, *5*, 61–116.
- Uhlig, J., Solga, H., & Schupp, J. (2009). Bildungsungleichheiten und blockierte Lernpotenziale: Welche Bedeutung hat die Persönlichkeitsstruktur für diesen Zusammenhang? [Inequality in Educational Opportunities: Underachievement and the Role of Personality Traits]. *Zeitschrift für Soziologie*, *38*(5), 418–440. <https://doi.org/10.1515/zfsocz-2009-0505>

- Uno, M., Mortimer, J. T., Kim, M., & Vuolo, M. (2010). "Holding on" or "coming to terms" with educational underachievement: A longitudinal study of ambition and attainment. In: S. Shulman & J.-E. Nurmi (Eds.), *The role of goals in navigating individual lives during emerging adulthood*. *New Directions for Child and Adolescent Development*, 130, 41–56. <https://doi.org/10.1002/cd.280>
- Veas, A., Castejón, J. L., O'Reilly, C., & Ziegler, A. (2018). Mediation analysis of the relationship between educational capital, learning capital, and underachievement among gifted secondary school students. *Journal for the Education of the Gifted*, 41(4), 369–385.
- Vlahovic-Stetic, V., Vidovic, V. V., & Arambasic, L. (1999). Motivational characteristics in mathematical achievement: A study of gifted high-achieving, gifted underachieving and non-gifted pupils. *High Ability Studies*, 10(1), 37–49. <https://doi.org/10.1080/1359813990100104>
- Weiß, R., & Osterland, J. (2012). *Grundintelligenztest Skala 1 (CFT 1, 6)* (überarb). Hogrefe.
- White, S. L., Graham, L. J., & Blaas, S. (2018). Why do we know so little about the factors associated with gifted underachievement? A systematic literature review. *Educational Research Review*, 24, 55–66. <https://doi.org/10.1016/j.edurev.2018.03.001>
- Woodward, L. J., & Fergusson, D. M. (2000). Childhood peer relationship problems and later risks of educational under-achievement and unemployment. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(2), 191–201. <https://doi.org/10.1017/S002196309900520X>
- Ziegler, A., Ziegler, A., & Stoeger, H. (2012). Shortcomings of the IQ-based construct of underachievement. *Roeper Review*, 34(2), 123–132. <https://doi.org/10.1080/02783193.2012.660726>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Alexander Dings is a research assistant at the individual differences and psychological diagnostics unit at Saarland University, Germany. His research interests include differences in academic achievement as well as mental illness stigma.

Frank M. Spinath is a full professor of psychology and head of the individual differences and psychological diagnostics unit at Saarland University, Germany. His research interests focus on the behavioral genetics of intelligence and personality, social inequality and educational success, as well as the utility and validity of novel test formats to measure cognitive ability.